

スペイン語音素の有標性と頻度. 辞書とテキスト

上田博人

内容

要旨	2
1. 序論	3
2. 分析	13
2.1. 音素	13
2.2. 音節	18
2.2.1. 音節構造	21
2.2.2. 音節の中の音素	25
2.2.3. 音素的音節	30
2.3. 語	33
3. 応用	36
3.1. 年代	37
3.2. 歴史	37
3.3. 地理・歴史	38
3.4. 社会	39
3.5. 考察	40
4. 結論	43
参考文献	46
【付録】	52

要旨

一般に言語類型と言語獲得に表れる音素の無標性・有標性の対立と出現頻度の関係について、有標項が無標項よりも低頻度であることが予想される。この予想がどれほど確かなものであるかを見るために、有標性を計測する有標度を設定し、スペイン語辞書の見出し語と偏りのないテキストにある語形をそれぞれ音素表記して、音素・音素的音節・音素的語の出現頻度を求め、それと質的に定位された有標性の相関を見る。辞書として Real Academia Española が発行した *Diccionario de la lengua española*. (23a ed., 2001) に収められた全見出し語 (87,515 語) の音素表記を使用する。書き言葉の資料として戯曲・小説・随筆・新聞・科学技術文を同じ割合で収録し、使用語の頻度を計算した Juilland and Chang-Rodriguez (1964) にある変化形の音素表記を使い、話し言葉の資料として Santander (España) に住む話者の性・年齢・学歴を考慮して選んだ 18 人の自由な会話の転写を用いる (Martínez y Ueda 2021)。

Marca y frecuencia de los fonemas españoles. Diccionario y textos.

Hiroto Ueda

Resumen:

En general, según la relación entre la marcación de los fonemas que aparecen en las distintas tipologías lingüísticas y la adquisición del lenguaje y la frecuencia de aparición, se espera que los fonemas marcados sean menos frecuentes que los no marcados. Para confirmar la validez de esta hipótesis, hemos establecido la fórmula del grado de marcación y hemos analizado con ella las frecuencias de aparición de los fonemas, las sílabas y las palabras fonológicas en los lemas recogidos en el diccionario (RAE 2001: *Diccionario de la lengua española*.), los textos escritos (Juilland and Chang-Rodriguez 1964) que recogen palabras de teatros, novelas, ensayos, periódicos y ciencias y los textos orales transcritos en Santander (España) de hablantes clasificados por sexo, edad y nivel de educación (Martínez y Ueda 2021).

1. 序論

Trubetzkoy (1939, 1980: 81) は音韻論研究の重要な主題の1つとして「有標・無標」の概念を次のように提起した。

<欠如的>対立. これは、対立項の一方が或る標識の存在によって、他方がその標識の欠如によって、特徴づけられるような対立である：たとえば、「有声」-「無声」，「鼻音化」-「非鼻音化」，「円唇」-「非円唇」など、標識の存在によって特徴づけられる対立項を<有標の>項といい、その標識の欠如によって特徴づけられる対立項を<無標の>項という。

そして「中和」が可能な欠如的音韻対立の中で、中和位置に現れる項が無標項に相当する(同: 88)。「自然な」無標性は、音声が発せられるとき、ふつうの呼吸を乱すことが最も少ない対立項に帰せられる(同: 107)。また、Jakobson and Waugh (1979: 1986: 95) は「有標項」の特色は特殊な文脈にしか現れないことである、と述べている。無標項が「自然に・ふつうに」出現するのに対し、有標項は「特殊な」出現の仕方をするのならば、無標項の頻度は高く有標項の頻度は低くなるはずである。田中(編)『現代言語学辞典』(1988: s.v. marked) は有標性を認める基準として頻度を含めていないが、Martinet (ed.)『言語学事典』(1965/1977: 221-223; 1969/1972: 341) は標識(無標・有標)の副次的な基準として頻度を認めている。そのとき、頻度の高さの不釣り合いが著しければ、中和がまったく起きていなくても、高頻度の項を無標の項として立てることが正しい、と述べている。

下の図-1a,b,c,d は、有標性を考慮しない二項対立(X, Y) (図-1a), 無標項(U)と有標項(M)を区別する二項対立 (図-1b), さらに有標項の少頻度を考慮する二項対立 (図-1c), そして最後に本研究が提示する頻度の連続性を考慮する多数項の考え方を示す (図-1d)。図-1dの有標項(m)はUから分離して周辺に位置する。

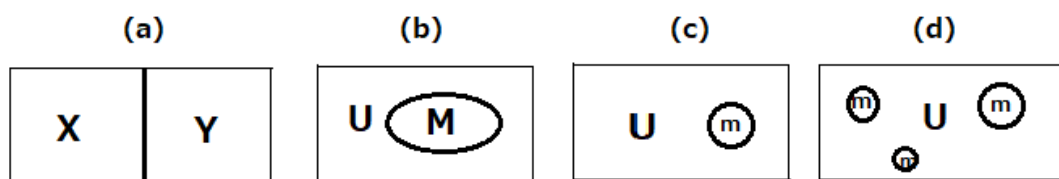


図-1a,b,c,d. (a) 二項対立(X, Y), (b) 無標項(U) / 有標項(M),
(c) 有標項=低頻度, (d) 連続的頻度を有する多数項

図-1a は「或る特徴の 2 つの段階とも、また或る特徴の存在または欠如ともみなせないような対立」(Trubetzkoy 同: 82) であり、たとえばドイツ語の p-t, f-k など、この対立はどの体系においても最も数が多い、という。当然スペイン語でも最も多い対立であるが、本研究では扱わない。注目すべきは図-1b が示す有標項の位置づけである。これは、デフォルトの無標項に対して、なんらかの標識が加えられて特別な有標項が成立する関係を示している。この有標性の概念は構造主義言語学、生成音韻論、さらに比較的最近の最適性理論でも中心的な基礎を成している。有標項の生起が少数であることを示す図-1c の考え方は言語単位の頻度の計測を重視するコーパス言語学の方法で取り組むことができる。最後に図-1d は多数の有標項が存在することを示している。有標項は異なり数 (type) は多いが、それぞれの延べ数 (token: 頻度) は低い。本研究では頻度の連続性を数量化することによって、有標性の連続的序列を設定する。

Zipf (1936) は、とくに有標性を考慮していないが、各種資料を基にして音素の複雑さ (たとえば無声音に対して有声音、無気音に対して有気音) が増すほどその頻度が低くなる、という原則を挙げている。そこで、ふつう複雑な音素は単純な音素よりも特殊であるので、それが有標の音素を指していると考えれば、やはり有標の音素は無標の音素よりも頻度が低い、ということになる。また、量的な観点から、たとえば、音節の頭位 (onset) にある /t/ (*ta.za*) と /tr/ (*tra.za*) を比べると /tr/ の方が複雑 (有標) であるので、その頻度が低いことが予想される¹。

有標性と頻度の関係を見るためには、世界の言語を比較することもあり、特定の言語の中で調査することもある (Rice 2007: 94)。田中 (2009: 18-23) は一般に (a) 言語類型と、(b) 言語獲得に表れる音素の無標性・有標性について次の表を掲げている²。

音素	特徴	無標	有標
母音	調音位置	低母音 (a), 高母音 (i, u)	中母音 (e, o, ε, ɔ)
子音	声帯振動	無声音 (p, t, k, f, s, ʃ)	有声音 (b, d, g, v, z, ʒ)
	調音位置	舌頂音 (t, d, s, z, ʃ, ʒ)	唇音 (p, b, f, v), 舌背音 (k, g)
	調音様式	閉鎖音 (p, b, t, d, k, g)	摩擦音 (f, v, s, z, ʃ, ʒ)

表-1. 無標の音素と有標の音素 (田中 2009: 21)

¹ Trubetzkoy (op.cit.: 278-279) は Zipf の理論に関して「或る欠如的対立の 2 つの項のうち、その無標項は連続する談話において有標項よりもヨリ頻繁に現れる」という定式は全体的に正しいであろうが、例外なしの法則と見なしてはならないと述べている。

² 同じ音素 (たとえば/d/) は声帯振動の特徴から見れば有標であるが、調音位置の特徴から見れば無標となる。

(a) 世界の言語の類型を知るためには、系統的に分類して選択された 317 言語の音素目録を整理した Maddieson (1984) の豊富な資料が役立つ。本研究の仮説を設定するときにこれを使用する。

(b) 言語獲得に関して、Jakobson (1968 : 47-51) は子供が最初に獲得する母音は a であり、調音様式は鼻音と口音 (oral) である、と言う (mama, papa)。次に調音点として唇音と歯音の対立を獲得する (papa - tata)。次に母音 i, u を獲得する。そして、子供が後で獲得する音素を、失語症患者は最初に失う、と述べている (同: 59-66)。

彼はさまざまな事例を示した後で次の結論に到達する：「子供の音韻の獲得と失語症患者の音声の障害は世界の言語の音韻目録と音韻史と連帯する法則に基づいている」(同: 92)。しかし、世界の言語の音韻についての言及はあるが、音韻史については扱っていない。そこで、Jakobson の論旨に沿って推測すれば、音韻変化によって新しく出現する音素は子供が新たに獲得する音素のように有標であり、音韻変化によって消失する音素は失語症患者が失う音素のようにこれも有標ということになる。そして、逆に、音韻変化が影響しない音素は無標になるはずである。たしかにラテン語とスペイン語の音韻を比較すると (Alarcos Llorach 1971 : 211, 217, 230, 240, 254, 265), それぞれの言語に排他的に存在する音素の有標性は高い (lat: 長母音, /h/, 語末の /-m/ など ; esp: /ye/, /we/, 硬口蓋子音 (/ç/, /j/, //), 中世スペイン語の有声摩擦音など)。次のリストの丸括弧 (...) は次の段階で失われた音素を示し、影をつけた部分は新しく生まれた音素を示す (Lat.: latín, His.: diasistema hispánico, Med: español medieval, Mod: español moderno)³。

- (1) a. Lat. a, (ā), e, (ē), i, (ī), o, (ō), u, (ū), p, t, k, b, d, g,
f, s, (h), m, n, r, l
b. His. a, (ɛ), e, i, (ɔ), o, u, p, t, t^s, k, b, d, g, (d^z), g,
f, s, (v), m, n, ñ, r, l, ʎ
c. Med. a, ye, e, i, we, o, u, p, t, (t^s), ç, k, b, d, g,
f, s, (ʃ), (v), z, (ʒ), m, n, ñ, r, ř, l, ʎ

³ Trask (200: 207) は有標性の歴史的変化 (markedness shift: 例: ing. *f-v* の単複 *leaf-leaves* が無標→有標に変化) を説明している。一方 Anderson (2001: 31) はたとえば異音の強化が一般に閉鎖音・舌頂音 (無標) より摩擦音・後舌音 (有標) より早く起きたなどを示している。両者とも有標性が歴史変化の「過程」に関係することを述べているが、ここでは、歴史的变化の「結果」として生まれる音素または失われる音素が有標項である、という仮説を立て、それに従えばスペイン語の高母音 /i, u/ は有標であることを論じる (後述)。なお Martinet (1955, 1964: 208-209) は頻度を音素結合の単純化の原因の 1 つとして考えられるが、その頻度は確実に例外的なものでなければならない、と述べている。

d. Mod. a, ye, e, i, we, o, u, p, t, č, k, b, d, g,
f, (θ), s, x, m, n, ñ, r, ř, l, (λ)

スペイン語話者の両親と英語話者の両親をもつ一歳児の発音を記録した Oller, D. K. and R. E. Eilers (2008) によれば両言語に共通する次の傾向を示した (>は頻度の大小関係を示す)。

- (2) 単子音>子音結合, 語頭子音>語末子音, 語頭: 閉鎖音>摩擦音・破擦音, 語末: 摩擦音・破擦音>閉鎖音, 語頭: 無気音>有気音, 語末: 無声音>有声音, 母音の前の滑脱音(半子音)>母音の後の滑脱音(半母音), 舌頂音>舌背音, 母音⁴: a > e > i > o > u.

上の頻度の大小関係はスペイン語音素の有標性を反映しているように見える。

次の表とグラフは Maddieson (1984: 205-262) が世界の言語を系統的に分類して選択し (317 言語), それぞれの音素をもつ言語数をまとめたものである。スペイン語が含まれているので, スペイン語音素と対応する音素をもつ言語の数とそのパーセント (母数 317) を算出した。

N	音素	言語数	%	N	音素	言語数	%	N	音素	言語数	%
1	n	315	99.4%	11	w	238	75.1%	21	r	80	25.2%
2	t	309	97.5%	12	b	198	62.5%	22	x	76	24.0%
3	m	299	94.3%	13	d	195	61.5%	23	ř	57	18.0%
4	k	283	89.3%	14	g	175	55.2%	24	θ	18	5.7%
5	s	276	87.1%	15	l	148	46.7%	25	λ	15	4.7%
6	a	274	86.4%	16	č	141	44.5%				
7	i	271	85.5%	17	f	135	42.6%				
8	y	271	85.5%	18	o	133	42.0%				
9	p	263	83.0%	19	e	113	35.6%				
10	u	254	80.1%	20	ñ	107	33.8%				

表-2. スペイン語音素と対応する音素をもつ言語の数とパーセント

⁴ 英語の異音をスペイン語の 5 音素に分類し集計した。

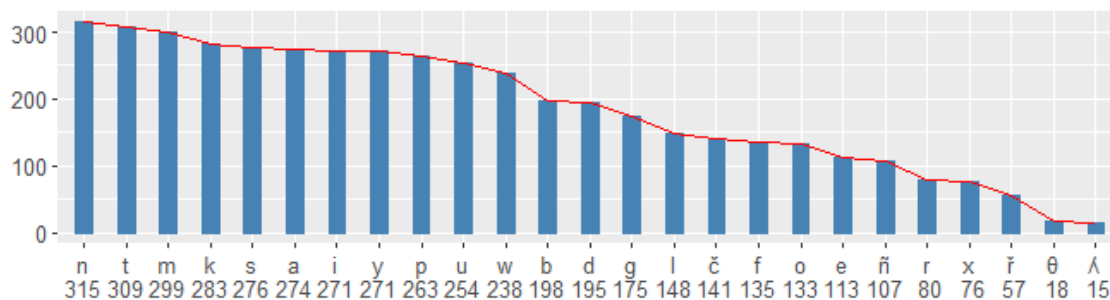


図-2. スペイン語音素と対応する音素をもつ言語の数⁵

上の表とグラフを見ると、スペイン語音素と対応する音素をもつ言語の数の序列がスペイン語音素の有標性の序列に従っているように見える。

Hyman (1975: 145-146) は普遍的有標性の特徴として、(a) 有標項は無標項に何らかの要素が付加された項である (例: /kʷ/ [有標]: /k/ [無標]), (b) 無標項は有標項よりも頻度が高い (よって高頻度の /a/ は無標の母音), (c) 中立性のある音素は無標である (例: 日本語の /u/), (d) 生産性・規則性がある音素は無標項である (例: 英語の 2 音節名詞の第 1 音節強勢配置)。この中で (a, c, f) は質的性質であり、(b) は量的性質なので観点が異なる。

実際に以下で見るように、質的に区別された無標の音素・構造の頻度が有標項の音素・構造の頻度よりも低いこともある (Trubetzkoy op.cit.: 277-280)。たとえばスペイン語の高母音 /i, u/ は言語類型では無標とされるが (表-1)、以下で見る頻度調査 (辞書とテキスト) では /e, o/ よりも低頻度で出現する。よって、頻度から見れば /a, e, o/ は無標であり、/i, u/ は有標となる。これはラテン語の弱勢の語末母音 /ī, i, ē, e/, /ū, u, ō, o/ がそれぞれスペイン語の /e/, /o/ に変化したことも根拠の 1 つとなる。

さらに、無標 /a, i, u/: 有標 /e, o/ という識別には次のような疑問点がある。たしかに世界の言語の中で 3 母音体系を有する言語には /a, i, u/ があり、/a, e, o/ ではない。そして、5 母音体系を有する言語には /a, e, i, o, u/ がある。この事実から 5 母音体系 /a, e, i, o, u/ の中の /e, o/ が有標である、とされている (表-1)。ここでは 3 母音体系の中の /i, u/ が 5 母音体系の中の /i, u/ に対応することを前提にしている (図-3.A)。しかし、2 つの体系を全体的に比較すれば、3 母音体系の中の /i, u/ は 5 母音体系の中の /i, u/ に対応するのではなく、/i~e, u~o/ に対応する、と見做すべきであろう⁶。そうすると 5 母音体系の /i~e, u~o/ の中で /e, o/ が無標の /a/ に

⁵ 本研究のグラフの作成には R のパッケージ ggplot2 と Excel を使用した。

⁶ 同じことは 5 母音体系の言語 (/a, e, i, o, u/) と 7 母音体系の言語 (/a, ε, e, i, ɔ, o, u/) についても言える。ここでは前者の /e, o/ が後者の /e, o/ では

近いので無標となり， /a/ から遠い /i, u/ が有標となるはずである（図-3.B）。下の図-3のAの見方では，言語Yの /e, o/ が言語Xに存在しないので特殊であるため有標とされる。一方，Bの見方では，言語Yの /i/: /e/, /u/: /o/ の対立が言語Xに存在しないのでそれぞれの対立（有標:無標）の中で /a/（無標）から最大の距離がある /i/, /u/を有標とする。Aに従って /e/, /o/ を有標とすると， /a/ - /e, o/ - /i, u/ という開口度の配列が無標-有標-無標という順番になり不自然である。無標を出発点にして有標に向かう推移を考慮するならば，無標項の位置は単一または連続する領域でなければならない。Bによれば /a/ - /e, o/ - /i, u/ という開口度の配列は無標から有標へ進む自然な段階的配列となる。本研究ではBの見方を採用する。

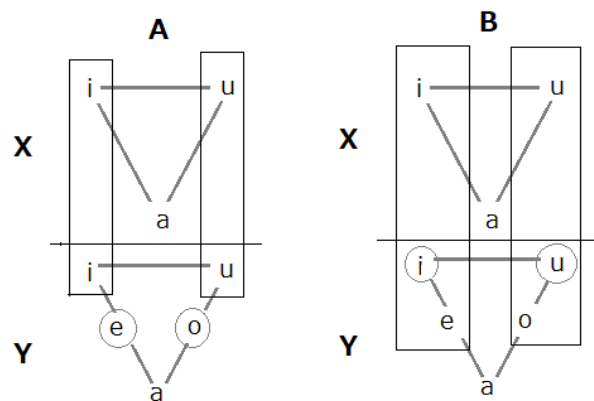


図-3a,b. X: 3 母音体系と Y: 5 母音体系の対応 A, B (○:有標)

母音の調音位置については言語獲得 (2)，世界の言語類型 (表-2, 図-2)，そして，スペイン語音韻変化 ('ε > 'ye, 'ɔ > 'we) (1) を考慮して，無標: 前母音 (e, i, y) » 有標: 後母音 (o, u, w) とする。母音の開口度を主対立，調音位置を副対立とすれば，無標→有標の序列は a » e » o » i » u » y » w となる。

子音の調音位置については，Schane (1973/1980: 147-148) の次の図式

*	p	t	č	k
anterior (前方音)	U	U	M	M
coronal (舌頂音)	M	U	M	U

表-3. 子音の調音位置 (Schane 1973/1980: 147-148)

に従って，無標 (U) » 有標 (M) の序列を次の順番とする。

なく， /ε ~ e, ɔ ~ o/ に対応する。しかし，この場合，母音 /a/ からより遠い，という理由だけで /e, o/ を有標とし， /ε, ɔ/ を無標とすることはできない。たとえば東アンダルシア方言の開母音化は音節末子音の氣息音化や脱落という有標な過程によって生じるので，開母音の方が有標となる。

(3) 歯・歯茎音 (UU) » 唇音 (UM)・軟口蓋音 (MU) » 硬口蓋音 (MM)

ここでは、子音の有標性の段階は口腔の中央左 (歯・歯茎音: 下図 2) から周辺 (唇音: 下図 1・軟口蓋音: 下図 4) に進み、最後は中央右 (硬口蓋音: 下図 3) となっていて、先の母音の有標性の段階 (図-3:B) と比べて不自然に見えるかもしれない。しかし、図-3:A の不自然さは無標 (a) - 有標 (e, o) - 無標 (i, u) という推移に見られ、ここでは無標項が序列の両側に現れている。一方、子音の調音位置では、無標項ではなく有標項が序列の両側 (下図 1, 4) に現れているので、母音の有標性の序列と本質的に違う。さらにの最大有標項である硬口蓋音の位置を含めて、先の Schane の図式に従えば、上の有標性の序列は納得できる。



図-4. X: 子音の調音位置 (○:有標項, ◎最大有標項)

調音様式について窪菌 (1998: 25-26) は

(4) 閉鎖音 » 破擦音 » 摩擦音 » 鼻音 » 側音 » 半母音

の順で有標性が増す、と言う。これを採用すれば表-1 の閉鎖音 (無標) » 摩擦音 (有標) の区別よりも範囲が拡大する。興味深いことに、これは音節を構成する原理となる聞こえ度 (sonority) の序列 (Harris 1983: 15 ; 後述 2.2) と同じである。Hooper (1976: 199) によれば、この序列の右の方がもっとも母音らしさ (vowel-like) を示すので、左の方が子音らしさを示すことになる。よって、左の方が子音として無標であり、右に進むほど有標性が高くなると考えられる。なお、スペイン語の r, ʎ を含めるために上の序列の中の側音 (lateral) を流音 (liquid) とする。

窪菌・本間 (2002: 107, 133) と Guffey (2002: 41) は音節構造との関係で、とくに阻害音 (obstruent: 閉鎖音・破擦音・摩擦音) と共鳴音 (sonorant: 鼻音・流音・半母音) の区別の重要性を強調している。この区別は、先の窪菌 (1998: 27) に従えば阻害音が無標、共鳴音が有標となる。

他に, 集約性(compact) » 拡散性(diffuse) ($s > \theta$), 中断性(interrupted) » 連続性(continuant) ($r > l$), 弛緩性(lax) » 緊張性(tense) ($r > \dot{r}$) を含める⁷。

以上の先行研究の成果を参照してスペイン語の音素の無標 : 有標の対立を仮説として次のように設定する。ここで, '»'の左を無標項, 右を有標項とする (無標 » 有標)⁸。不等号記号 '>' の左は右よりも頻度が高い, と予想する。

仮説-I : スペイン語音素の有標性の序列 (無標»有標) [対立数: 46]

- (5) a. 母音/半母音 : $a \gg e \gg o \gg i \gg u \gg y \gg w$:
 $a > (e, o, i, u, y, w)$; $e > (o, i, y)$; $o > (u, w)$; $i > (u, y)$; $u > w$
 [対立数: 14].
- b. 声帯振動 : 無声子音 » 有声子音: $p > b$; $t > d$; $k > g$ [対立数: 3].
- c. 調音位置 : 歯音/歯茎音 » 唇音/軟口蓋音 » 硬口蓋音
 $t > (p, k) > \check{c}$; $d > (b, g)$; $(\theta, s) > (f, x)$; $n > (m, \tilde{n})$; $m > \tilde{n}$; $l > \Lambda$
 [対立数⁹: 15].
- d. 調音様式 : 閉鎖音 » 破擦音 » 摩擦音 » 鼻音 » 流音:
 $p > f$; $t > (\theta, s)$; $k > x$; $b > m$; $d > (n, r, l)$; $n > (l, r, \dot{r})$ [対立数¹⁰: 11].
- e. 集約性 » 拡散性 : $s > \theta$ [対立数: 1].
- f. 中断性 » 連続性 : $r > l$ [対立数: 1].
- g. 弛緩性 » 緊張性 : $r > \dot{r}$ [対立数: 1].
- h. 少数 » 多数 : $t > tr$; $CV > CVC > CVCC$ など (後述 2.2.1).
- i. 一般 » 特殊 : $te.'nis.ta > 'Fr\acute{o}.mis.ta$ など (後述 2.1).

「対立数」はすべての可能な関連する「最小対立」の数の和を示す。「最小対立」とは声帯振動・調音位置・調音様式 (母音の開口度を含む) のいずれか 1 つだけが異なる対立を指す。よって, たとえば $e-u$ は開口度と調音位置が異なるので最小対立ではない。

音素, 音節, 音素的語の出現頻度を使ってその有標性が求められれば, その有標性の程度, すなわち「有標度」が計算できる。たとえば, 開音節と閉音節を比べて表-4 のようなデータが得られたときの有標度 (GM:

⁷ Jacobson, Fant and Halle (1952/1965 : 44-48, 35-38, 60-64), Alarcos Llorach (1971: 179).

⁸ 無標・有標の区別は範列的な音韻対立の中で認められるので, 連辞的に対比する母音と子音や母音と半母音の間では区別されない。よって, 母音どうし, 子音どうしの間で区別される。そして, 子音は声帯振動・調音位置・調音様式の 1 点だけが異なる対 (ペア) で区別される。

⁹ $t-p, t-k, t-\check{c}, p-\check{c}, k-\check{c}, d-b, d-g, \theta-f, \theta-x, s-f, s-x, n-m, n-\tilde{n}, m-\tilde{n}, l-\Lambda$: 計 15 個。

¹⁰ $p-f, t-\theta, t-s, k-x, b-m, d-n, d-r, d-l, n-l, n-r, n-\dot{r}$: 計 11 個。

Grado de marcación) は次のようになる (後述 2.2.1)。

種類	頻度	相対頻度	有標度 (GM)
開音節	232 850	0.718	0.000
閉音節	91 344	0.282	0.608

表-4. 開音節と閉音節 頻度と有標度

次が本研究で仮説-IIとして提案し、その有用性を検証する「有標度」の式である。

仮説-II: 有標度 (Grado de marcación: GM, F: 頻度, Max: 最大頻度):

$$(6) \quad GM = 1 - \frac{F}{Max}$$

たとえば、先のデータ (表-4) の閉音節の有標度は次のように計算される。

$$GM (\text{表-4: 閉音節}) = 1 - \frac{91344}{232850} = 0.608$$

有標度 (GM) は有標性が最大するとき、すなわち頻度 (F) がゼロ (0) のとき、最大の 1 になる。逆に、無標項の有標度はゼロ (0) になる。よって、有標度は最小値 (0) と最大値 (1.000) との間に様々な段階を連続的に表示する¹¹。

有標性の程度を見るために、有標度ではなく、単純に相対頻度 (またはパーセント) を使用することも考えられるが、総和と比較する相対頻度は無標項と比較しているのではないため、それが有標性を表すことにはならない。一方、有標度は無標項に対応する最大頻度を母数として使っているため、有標性を示す指標として適している¹²。また、相対頻度は総和を母数とするため、微小な数値になることが多い¹³。そして、相対頻度は有標性と大小関係が逆になる (相対頻度が小さいほど有標性は高くなる)¹⁴。

本研究では、各種の語彙資料を使ってスペイン語の音素・音節の頻度と有標性の関係を調べ、先に挙げた仮説-I「スペイン語音素の有標性の序列」(5)の有効性と、仮説-II「有標度」(6)の有用性を検証する。資料は

¹¹ ただし、有標度の最大値 (1.000) は音素の頻度 (F) がゼロ (0) のときの値 (存在しない音素の有標度) なので現実的ではない。

¹² GM の第 2 項 (F/Max) は「無標性」の度合を示す。

¹³ 乗数を 1000 にしたパーミルや、さらに大きな乗数を使うことも考えられる。

¹⁴ 相対頻度の 1 の補数を使うことも考えられる (1 - 頻度 / 総和)。

Trubetzkoy (同: 282-285) に従って、辞書とテキストを使う。大規模な辞書は言語のほとんどすべての語彙を載せているので、形態の可能性をほぼ網羅しているが、代表形（見出し語）に限り、名詞・形容詞・動詞などの変化形は載せていない。一方、テキストは実際の使用形（変化形）を載せているが、言語の語彙を網羅していない。よって、辞書は音韻の実現の理論的可能性を示し、テキストの中の音素の頻度はその実際の蓋然性を示していることになり、両者は相補的な関係になる。また、テキストの中の音素の頻度は冠詞や前置詞などの高頻度語に使用される音素の頻度が反映されるが¹⁵、辞書の資料にはそのような影響はない。辞書として、Real Academia Española が発行した *Diccionario de la lengua española*. (DLE, 23a ed.) に収められた全見出し語 (87,515 語) の音素表記を使用し¹⁶、書き言葉のテキストとして戯曲・小説・随筆・新聞・科学技術文を同じ割合で収録し (各 10 万語)、使用語の頻度を計算した Juilland and Chang-Rodríguez (1964) にある全変化形の音素表記を用いる。話し言葉のテキストとしてスペインの Santander に住む話者の中から性・年齢・学歴を考慮して選んだ 18 人の会話の転写を用いる (Martínez y Ueda 2021)¹⁷。

¹⁵ テキストの中の音素の頻度は確かに少数の異なり語数をもつ高頻度語の音素の頻度が反映するが、一方、低頻度語の異なり語数は多数を占めるので低頻度語の音素の頻度も十分に反映する。よって、テキストの中の音素の頻度数調査はテキストの実態を示しているものとして有効である。

¹⁶ Domínguez (2021/08/26) の資料を使用する。ただし、*b, c, nn* などの文字名や略語は除外した。*psicología* や *mnemotecnia* などの特殊な文字列は一般の発音 (/s/, /n/) にした。一部の転記ミスを修正した。

¹⁷ Santander では一般に /s/ : /θ/ と /l/ : /ʎ/ の対立を保持している。3つの資料を比較するために、辞書と書き言葉資料も同じ対立を用いる。このようにスペイン語の守旧的な方言を用いて資料を作成すれば、必要に応じて seseo (/θ/ → /s/), ceceo (/s/ → /θ/), yeísmo (/ʎ/ → /l/) のある方言をベースにした辞書と書き言葉資料の音素の頻度の計算ができる (対立する音素の頻度を合計すればよい)。逆に、たとえば seseo の方言で計算した音素の頻度から、/s/ : /θ/ を区別する方言をベースにした資料にある音素の頻度の計算はできない。よって、守旧的な方言は全体的・基盤的な音素体系を有する、と言える。

2. 分析

以下では音素→音素的音節→音素的語という順番に従って音素の頻度を計測する。すべての音素の頻度は基本的に音素的音節と音素的語の中の音素の頻度に反映される，と予想されるため，音素の頻度だけを考察すれば十分である，とも考えられる。しかし，それぞれの音素の分布を見なければ，音韻の構造と体系が理解できない。また，以下で見るように，音節内の各位置での音素の頻度分布は全体の頻度分布とは大きく異なっている部分がある。そこで，音素の環境として重要な音節の構造を考慮した。そして，音節と語の規模（長さ）も音韻の有標性を示す指標となる¹⁸。

2.1. 音素

次は辞書 (D)，書き言葉テキスト (E)，話し言葉テキスト (O) のすべての語形に現れた音素の頻度である¹⁹。

¹⁸ 実際には語内の音素，語内の音節，強勢の有無なども詳しく観察すべきであるが，これらについては稿を改めなければならない。

¹⁹ 以下の統計処理では R を使用した (R Core Team 2021)。

D	音素	頻度	有標度	E	音素	頻度	有標度	O	音素	頻度	有標度
1	a	109 112	0.000	1	e	282 076	0.000	1	e	91 451	0.000
2	o	76 028	0.303	2	a	235 526	0.165	2	a	75 847	0.171
3	e	73 948	0.322	3	o	180 411	0.360	3	o	64 871	0.291
4	r	58 080	0.468	4	s	156 909	0.444	4	s	51 253	0.440
5	n	51 298	0.530	5	n	139 570	0.505	5	n	45 738	0.500
6	i	49 736	0.544	6	d	105 999	0.624	6	r	30 458	0.667
7	t	41 624	0.619	7	r	105 168	0.627	7	k	27 921	0.695
8	s	33 451	0.693	8	l	100 394	0.644	8	i	27 887	0.695
9	k	31 177	0.714	9	t	80 637	0.714	9	t	26 852	0.706
10	l	29 664	0.728	10	i	79 890	0.717	10	d	24 655	0.730
11	d	28 771	0.736	11	k	74 804	0.735	11	l	24 059	0.737
12	b	21 349	0.804	12	y	59 907	0.788	12	m	21 687	0.763
13	m	20 727	0.810	13	p	48 025	0.830	13	y	21 290	0.767
14	p	18 823	0.827	14	m	45 960	0.837	14	p	16 879	0.815
15	y	17 859	0.836	15	b	44 876	0.841	15	b	16 433	0.820
16	θ	17 762	0.837	16	u	44 352	0.843	16	u	13 597	0.851
17	u	15 398	0.859	17	θ	30 874	0.891	17	w	9 025	0.901
18	g	10 766	0.901	18	w	17 011	0.940	18	θ	8 157	0.911
19	ř	8 376	0.923	19	g	15 570	0.945	19	g	6 031	0.934
20	f	8 275	0.924	20	f	11 828	0.958	20	x	4 214	0.954
21	x	7 293	0.933	21	x	10 990	0.961	21	f	2 966	0.968
22	w	5 565	0.949	22	ř	10 612	0.962	22	č	2 143	0.977
23	č	4 774	0.956	23	ǻ	6 382	0.977	23	ř	2 047	0.978
24	ǻ	3 871	0.965	24	ñ	4 323	0.985	24	ǻ	1 740	0.981
25	ñ	2 022	0.981	25	č	4 122	0.985	25	ñ	1 240	0.986

表-5a,b,c. 音素の頻度. (a.[D]iccionario, b.[E]scrito, c.[O]ral)

この表を見ると次のことが確認できる。下のリストの中の「>」の記号の左が右よりも多いことを示す。たとえば a > o は a は o よりも頻度が高いことを示す。アステリスク (*) は例外を示す (D: 辞書; E: 書き言葉; O: 話し言葉)。

- (7)
- a. 母音/半母音 : a » e » o » i » u » y » w: 例外 *e > a (E, O)
 - b. 声帯振動 : 無声子音 » 有声子音: 例外 *b > p (D), *d > t (E)
 - c. 調音位置 : 歯音/歯茎音 » 唇音/軟口蓋音 » 硬口蓋音
例外 *o > e (D), *k > t (O)

- d. 調音様式：閉鎖音 » 破擦音 » 摩擦音 » 鼻音 » 流音：
例外 *m > b (E, O), *n > d (D, E, O), *r > d (D, O)
- e. 集約性 » 拡散性：例外なし.
- f. 中断性 » 連続性：例外なし.
- g. 弛緩性 » 緊張性：例外なし.

このように (7) は多少の例外はあるが (*で示した 12 個の対立), ほとんどが仮説-I (5) の有標性の序列通りであることを示している。この有標性と頻度順位の間多くの一致は偶然ではないと思われるので, 無標→有標と高頻度→低頻度の相関関係を否定することはできない。

例外はそれぞれの資料の中で有標度が比較的低い音素で起きている：e (有標度[E]: .000, 有標度[O]: .000), a (.165, .171); b (.804), p (.827); d (.624), t (.714); o (.303), e (.322); k (.695), t (.706); m (.837), b (.841); n (.530, .505, .500), d (.736, .624, .730); r (.468, .667), d (.544, .730)。有標度が.900を超える音素では例外はない。よって, 音素対立の多くは仮説-I (5) に従い, とくに有標性が高い音素は完全に仮説に従っている, と言える。この理由は有標度が高い音素は有標性 (出現頻度) を鮮明に現わし, 一方, 有標度が低い音素はその有標性が曖昧になることがあるためである, と考えられる。

有標度は連続する数値なので, 無標→有標という有標性の推移も連続する。よって, 二項対立を除けば質的に無標：有標の区分することはできない。仮に量的基準として有標度=.500で区分すれば, 辞書 (D) では a, o, e, r が無標項となり (表-5, 太字), その他が有標項となる。そのときの無標項は少数の高頻度音素であり, 有標項は多数の低頻度音素である。

次の図-5 は辞書 (Diccionario), 書き言葉 (Escrito), 話し言葉 (Oral) の音素使用頻度の順位を比較している (順位のをを使って並べ替えた)。

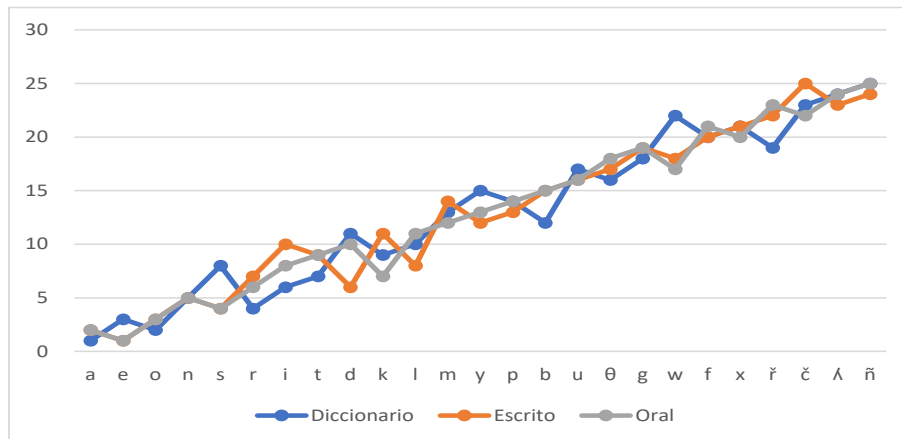


図-5. 音素使用頻度の順位 (辞書 / 書き言葉 / 話し言葉)

この図が示すようにテキスト（書き言葉 / 話し言葉）の /s/ が頻度がやや高い（順位の数値が低い）が、テキストでは名詞・形容詞などの複数形や動詞の2人称形などが頻用されるためであろう。書き言葉の /d/ の高頻度は最大頻度語の前置詞 *de* によるが、辞書ではそれが1回しかカウントされないため順位は落ちている。このように、資料の性質の違いによって若干の違いがあるが、全体的に見れば3つの資料の順位は類似している。

次の表-6は González Rátiva y Mejía Escobar (2011) による先行文献の資料と Navarro Tomás (1966), および本研究の結果（表-5）から頻度順位を比較したものである。末尾に順位の平均 (Me) と標準偏差 (SD) を載せた²⁰。

*	AL	GB	GG	LM	M.E	M.O	MG	N	P	PC	QE	R	U.D	U.E	U.O	ZR	Me	SD
e	2	1	1	1	2	1	2	2	1	1	1	1	3	1	1	2	1.4	0.6
a	1	2	2	2	1	2	1	1	2	2	2	2	1	2	2	1	1.6	0.5
o	3	3	3	3	3	3	3	3	4	4	3	3	2	3	3	3	3.1	0.4
s	5	4	4	4	5	4	7	4	3	3	4	4	8	4	4	4	4.4	1.3
i	4	6	6	6	4	5	4	10	6	5	5	5	6	10	8	10	6.3	2.0
n	11	5	5	5	6	6	5	5	5	6	13	15	5	5	5	5	6.7	3.1
r	6	7	7	8	7	7	6	6	7	7	6	14	4	7	6	6	6.9	2.0
t	8	8	8	7	10	8	8	9	9	10	7	8	7	9	9	9	8.4	0.9
l	7	10	11	9	8	9	13	7	8	9	9	6	10	8	11	7	8.9	1.8
d	9	11	10	11	9	11	9	8	10	8	8	7	11	6	10	8	9.1	1.5
k	10	9	9	10	11	10	12	11	11	11	10	9	9	11	7	11	10.1	1.2
m	13	13	12	12	13	12	10	12	13	13	12	13	13	14	12	13	12.5	0.9
b	12	15	13	13	15	15	11	15	15	15	15	11	12	15	15	12	13.7	1.6
p	14	14	15	14	14	14	15	13	14	14	14	12	14	13	14	14	13.9	0.7
u	15	12	14	15			14	17	12	12	11	10	17	16	16	16	14.1	2.2
θ	16			16	16	16	23	16			17	16	16	17	18	17	17.0	1.9
g	17	16	16	17	17	17	16	18	16	16	18	17	18	19	19	19	17.3	1.1

²⁰ それぞれの研究の音素分析法、とくに半母音と中和現象の解釈 (Hara 1973) が異なるので厳密に考えれば本来比較できないのだが、可能な限り二重母音は半母音と母音に分割し、/N/, /L/の中和音素はそれぞれ/n/, /l/として計算した。(AL: Alarcos Llorach (1961). Tipo de texto=Escrito; GB: Guirao y Borzone (1972), Escrito; GG: Guirao y García Jurado (1993), Oral; LM: Listerri y Mariño (1993); M.Escrito: Moreto et al. (1975), Escrito; M.Oral: Moreto et al. (1975), Oral; MG: Mejía-González (2008), Escrito; N: Navarro () Escrito; P: Pérez (2003); PC: Pineda/Cuétara (2004); QE: Quilis y Esgueva (1980); R: Rojo (1991), Escrito; U.D: Ueda (本研究), Diccionario; U.E: Ueda (本研究), Escrito; U.O: Ueda (本研究), Oral; ZR: Zipf y Rogers (1939), Escrito.)

y	22	19	17	24				19	14	19	20	21	22	15	12	13	15	18.0	3.6
w								19						22	18	17		19.0	1.9
f	18	17	19	20	19	20	20	21	17	17	20	19	20	20	20	21	20	19.3	1.3
x	19	18	18	18	20	19	18	24	18	18	19	18	21	21	20	22		19.4	1.7
ř	20	20	21	22	18	21	17	20	20	19	16	20	19	22	23	18		19.8	1.8
ʌ	21			19	21	18		22			22		24	23	24	21		21.5	1.9
č	23	21	20	21	23	22	21	25	21	21	23	21	23	25	22	24		22.3	1.5
ň	24	22	22	23	22	23	22	23	22	22	24	23	25	24	25	23		23.1	1.0

表-6. 音素の頻度順位 (先行文献と本研究の比較)

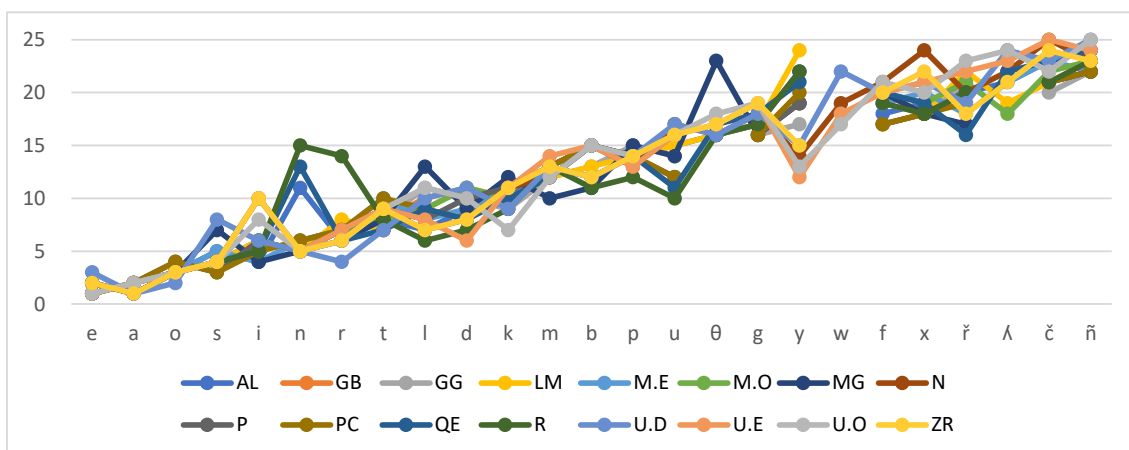


図-6. 音素の頻度順位 (先行文献と本研究の比較)

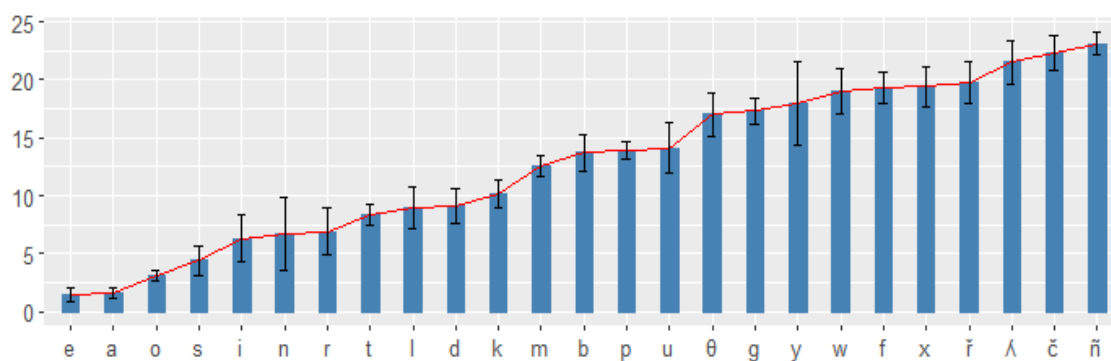


図-7. 音素の頻度順位の平均と標準偏差 (先行文献と本研究)

上の表-6と図-6,7を見ると一般に順位の傾向がよく類似し、順位の平均からの偏差は比較的小さいことがわかる。そして、全体の順位の推移の多くは仮説-I (5) の有標性の序列に従っている。例外は *e > a, *s > t, *n > d, *m > b, *b > p, *l > d であるが、これらの平均順位の偏差 (標準偏差) の範囲内にあるので決定的な反例ではない。資料によって逆転する可能性があるからである。また、例外のペアの順位の差が小さく順位が隣接するケー

が多い。よって、全体的に見ればグラフの左から右に移行するに従って有標性が増加する傾向がある、と言える。

2.2. 音節

スペイン語の音節を Harris (1983: 6-18) は Real Academia Española. (1973), Gil Fernández (1990), Martínez Celdrán (1994) のように Onset: (子音), Nucleus: (半子音・母音・半母音), Coda: (子音) という 3 要素 (8a) に区分するのではなく, (8b) のように O(nset)と R(hyme) という 2 要素に区分する方法を提案している²¹。

- (8) a. [Onset] - [Nucleus] - [Coda]: [b]_O [ie]_N [n]_C
b. [Onset] - [Rhyme (Nucleus - Coda)]: [b]_O [ien]_R

その理由として次の 2 点を挙げている。どちらも Rhyme という単位がなければ説明できないことになる。

- (9) a. 音節は *claus.tro*, *clien.te* などのように最大 5 要素を許容するが、同じ 5 要素であっても **muers.to* を許容しない。その理由は Rhyme (*uers*) の最大許容量が 3 要素だからである。
b. 単語末尾から 2 番目の音節の Rhyme が分岐しているときは (**te.lé.fos.no*, **te.lé.boi.na*), 後ろから 3 番目の音節に強勢は配置されない。

たしかに, (9a) **muers.to* にはスペイン語らしさがないが, 辞書 (DLE) には *eins.te.nio*, *bioin.for.má.ti.co*, *carst*, *ma.croins.truc.ción* などのように 4 要素の Rhyme を含む語が載せられている。また, 筆者には *clien.te* のほうが, *claus.tro* よりも自然さが感じられる²²。一方, *esn* などはほとんど不可能である。よって, 許容度の序列は次のようになる。

- (10) *clien.te* » *claus.tro* » *eins.te.nio* » **esn.co*

したがって, *muers.to* をスペイン語の完全に不可能な形態として排除するのではなく, その可能性 (posibilidad) はあるが蓋然性 (probabilidad) が少ない, と見るべきである。可能性があるという理由は, 'ue' も 'rs' も 2

²¹ 一方, D'Intrino, Teso y Weston (1995: 400-402) は [Onset - Nucleus] - [Coda] という 2 分法を採る。

²² 後述するように, 尾位が n である音節の数は膨大である (有標度が低い) が, ws という尾位の 2 音素をもつ音節の数は極めて少数である (有標度が高い)。

音素なので仮説-I (5h) に従えば有標でありながら、可能な連続であり (*puerta, persepectiva*), 'uers' は後述する sonority scale に合致する連続であるからである。それを困難にしているのは2つの有標要素が組み合わせられて極度の有標性 (二重の有標性) を示していることである、と考えられる。

(9b) はラテン語の強勢配置規則に遡るが²³, 現代スペイン語について Pensado (1985/2000: 489) は反例として *Frómista* のようなスペイン語地名や *Livingstone, Winchester* などの外国人名・地名を挙げている²⁴。これらは特殊な例であるが、一般的な例として Baquero (1999: 135) は *traigámoslo, recíbanlo, bájense* などの動詞活用形+代名詞の例を挙げる。たしかに, Harris (同: 10) は動詞形態を最初に除外しているが、この除外が規則の一般性を損ねていることも確かである。ここでも筆者は (9b) について確かに蓋然性が少ないが可能性は存在する、と考える。ただし極めて頻度が低く有標性が高い (5i)。

Harris (同: 14) は次のスペイン語音節の Rhyme type を挙げている²⁵。

- (11) a. V, VG, VL, VN, VO (+ s)
b. GV, GVG, GVL, GVN, GVO (+ s)

しかし、これでは *aun.que* (VGN), *trein.ta* (VGN), *vein.te* (VGN), *ist.mo* (VOO) などの高頻度語が除外されてしまう。これらは有標のスペイン語音節であるが、例外にしてはならない。

最適性理論 (Optimality theory: Prince and Smolensky 1993/2002) の枠組みに従って Colina (2009: 18-36) は音節に (12a) の言語普遍的な「聞こえ度の輪郭」(sonority contour) を設定している。これに基づいて頭位 (onset), 音節核 (nucleus), 尾位 (coda) でそれぞれ (12b, 12c, 12d) の制約の優先順位を示す。(12a) は連辞的な繋がりを示し, (12b-d) は範列的な選択の関係となる (後述 2.2.2)。(12b-d) は「有標性」('marcadez': Colina 2014: 86) の序列を示している (有標 >> 無標)。

²³ 出口 (1997: 35) が挙げている *ápache, *ápañe, *ápalle, *ápaje, *áparre などそれぞれ ch, ñ, ll, j, rr がラテン語の /子音.子音/ という連続に遡るので、語末から2番目の音節が閉音節となり, *te.lé.fos.no, *te.lé.boi.na と同類である。

²⁴ *Frómista* はゲルマン語の *fromists* 'primero' の女性形 *fromista* に由来する (Nieto Ballester, 1997: 169)。Alberto, Guillermo, Ricardo, Rodolfo などの人名もゲルマン語起源である (Albaigés Olivart, 1989)。これらはゲルマン語の特徴である語頭の強勢を持っていたが、後にこれらの人名ではラテン語の強勢配置に従った。地名の *Frómista* はゲルマン語の語頭の強勢を保持した。

²⁵ [V]owel, [G]lide, [L]iquid, [N]asal, [O]bstruent (Stop, Fricative).

- (12) a. 阻害音 < 鼻音 < 流音 < 滑脱音 < 母音
 b. 頭位： 母音 >> 滑脱音 >> 流音 >> 鼻音 >> 阻害音
 c. 音節核：阻害音 >> 鼻音 >> 流音 >> 滑脱音 >> 母音
 d. 尾位： 阻害音 >> 鼻音 >> 流音 >> 滑脱音 >> 母音

優先順位にある「>>」の記号の左側は右側より大きな違反となる。よって、阻害音は頭位で最も優先され、音節核と尾位では最も回避される。

(12b-d) は音声変化の出力の優先度を示すが、音声変化をしないときは入力の様式がそのまま忠実に守られる。その忠実性の制約の中に「削除をしない」という制約 (MAX-IO) と、「挿入をしない」という制約 (DEP-IO) がある。ここで「尾位に阻害音 (obstruent) が回避される」という有標性制約 (*CODA/obstruent) を使って次のような「タブロー」(tableau) ができる (Colina (2009: 29-30)。例は /obsoleto/ [osoletto] である (一部改変)。

	*CODA/obstruent	DEP-IO	MAX-IO
a. o .so.le.to			*
b. ob.so.le.to	*!		
c. o.be.so.le.to		*!	

表-7. /o.so.le.to/

ここでアステリスク (*) は、a. o.so.le.to, b. ob.so.le.to, c. o.be.so.le.to がそれぞれ MAX-IO, *CODA/obstruent, DEP-IO という制約に違反していることを示している。エクスクラメーションマーク (!) は重大な違反 (表の左側) を示し出力しないことを示す。出力されるのは軽微な違反 (MAX-IO) をした a. [o.so.le.to] ということになる (~~o~~で示す)。

丁寧な発音では音節末の/b/が出力されるので、制約の優先順位が変わって次のようなタブローになる。

	MAX-IO	DEP-IO	*CODA/obstruent
a. o.so.le.to	*!		
b. o ob.so.le.to			*
c. o.be.so.le.to		*!	

表-8. /ob.so.le.to/

Colina (同) が提示する最適性理論の枠組みでは音素の頻度は考慮されていないが、先述した頻度に基づく有標度 (GM) を導入すればタブローは

次のようになる²⁶。

	MAX-IO	DEP-IO	*CODA/obs. (GM)
a. o.so.le.to	*!		
b. o ob.so.le.to			* (.985)
c. o.be.so.le.to		*!	

表-9. /ob.so.le.to/ + 有標度 (GM)

GM=.985 は ob.so.le.to の有標性がかなり高いことを示している。よって、このタブローの *CODA/obs. という有標性制限は忠実性制限 (MAX-IO, DEP-IO) を超えて左の位置に移動して o.so.le.to が選択される可能性がある。このような観察は最適性理論が示す一般的なタブローでは不可能である。

本研究では言語普遍的な聞こえ度の輪郭 (12a) と有標性の序列 (12b-d) を念頭に置きながら、スペイン語固有の音節構造を設定し、聞こえ度の連辞的繋がりとその範列的な有標性の序列を観察し、同時に両者の間にある類似性の実態を探る。

以上の考察に従って、以下の記述では、すべての可能な音節を含む (13a) のような基本的な音節構造を設定して分析する。

- (13) a. (Cx) - V - (Cy)
 b. (C) - (C) - (C) - V - (C) - (C) - (C)

(13a) の V を単母音とし、Cx, Cy は 0 個以上の子音 (O, N, L, G) とし (13b), Cx, V, Cy に現れる音素の頻度を調査する。一般に、母音の前の滑脱音 (G) は母音 (V) と共に核の一部とし、母音の後の滑脱音 (G) は母音 (V) と離して尾位 (Coda) に含めるという方法が採られるが (Colina 2021: 141; Kaisse 2021: 151), ここでは、Hooper (1976) に従って O-N-L-G-V の連続性を重視し、直線的な構造 (13b) を設定する (上田 1975)。

2.2.1. 音節構造

はじめに、辞書と書き言葉の資料の音素表記を C と V から成る音節に変換し、開音節と閉音節の割合を見る。下の表によれば、一般に言われているように (Malmberg 1965, Granda Gutiérrez 1966), 確かに開音節が優勢である。しかし、閉音節もかなり現れていて、その有標度 (.606, .465) はそ

²⁶ GM = 1 - CODA/obs / CV = 1 - 2823 / 189609 = .985

れほど高くない²⁷。

辞書	頻度	%	有標度	書き言葉	頻度	%	有標度
開音節	232 850	71.8%	0.000	開音節	535 589	65.1%	0.000
閉音節	91 344	28.2%	0.608	閉音節	286 666	34.9%	0.465

表-10a, b. 開音節・閉音節 (a. 辞書 / b. 書き言葉)

表-10 で有標となる閉音節は、実は形態的に開音節と完全に離反して対立しているのではなく (14a), 開音節に子音が付加されている (14b), という事実に注目したい。

- (14) a. $CxV \leftrightarrow CxVCy$
 b. $CxV (Cy) : CxVCy$
 c. $CxV(Cy) \rightarrow CxVCy$

ここで、 $CxV \leftrightarrow CxVCy$ という選択的な関係 (二項対立) ではなく (14a), $CxV (Cy) : CxVCy$ という関係, つまり $CxV (Cy)$ には語尾の子音 Cy がない (無標), という欠如的關係を見るべきである (14b)。さらに, この無標:有標の間に「無標 \rightarrow 有標」という連続的な関係を捉えるべきである (14c)。

²⁷ Malmberg への反論として Mosonyi (1970), Catalán (1971), Torreblanca (1980) が重要である。

次がすべての音節パタンの出現頻度である (降順で並べ替えた)²⁸。

D	ED	頻度	GM	E	ED	頻度	GM	O	ED	頻度	GM
CV	0	189 123	0.000	CV	0	414 733	0.000	CV	0	142 704	0.000
CVC	1	63 265	0.665	CVC	1	170 483	0.589	CVC	1	50 473	0.646
CCV	1	25 252	0.866	VC	2	76 907	0.815	V	1	29 994	0.790
V	1	18 031	0.905	V	1	65 955	0.841	CCV	1	17 055	0.880
VC	2	15 012	0.921	CCV	1	54 195	0.869	VC	2	22 032	0.846
CCVC	2	11 278	0.940	CCVC	2	34 785	0.916	CCVC	2	10 920	0.923
CVCC	2	707	0.996	VCC	2	2 468	0.994	CVCC	2	559	0.996
VCC	2	670	0.996	CVCC	2	1 410	0.997	VCC	2	284	0.998
CCCV	2	416	0.998	CCCV	2	706	0.998	CCCV	2	98	0.999
CCVCC	3	246	0.999	CCVC	3	367	0.999	CCVCC	3	63	1.000
CCCVC	3	184	0.999	CCVCC	3	245	0.999	CCCVC	3	28	1.000
CVCCC	3	6	1.000	CCCVC	4	1	1.000				
CCVCCC	4	5	1.000								
VCCC	3	2	1.000								
CCCVC	4	1	1.000								

表-11a,b,c. 音節パターン・編集距離 (ED)・頻度・有標度 (GM)

(a: D.辞書 / b: E.書き言葉 / c: O.書き言葉)

どの資料でも音節パタンの出現個数のトップは CV であり、これが最も普通 (無標) の音節パターンである²⁹。第 2 位以下は 2.CVC, 3.CCV, 4. V (D); 2.CVC, 3.CCV, 4. V (E); 2.CVC, 3.V, 4. CCV (O) と続く。これらは 1.CV をベースにして、1つの要素の付加 (15a, b), または削除 (15c) で導かれる。

- (15) a. 1. CV _ + C → 2. CVC
 b. 1. _ CV + C → 3. CCV
 c. 1. CV - C → 4. V

VC は CV の語尾に C を付加して CVC とし、次に CV の語頭の C を削除して導く (16a)。または CV の語頭の C を削除して V とし、その後で語尾

²⁸ 音節末の 2 子音連続, 3 子音連続の例: CC: /bs/: subs.tan.cia, /ks/: sex.to, /ms/: háms.ter, /ns/: ins.ta.lar, /rs/: pers.pec.ti.vo; CCC: /ngs/: ángs.trom, /rps/: corps, /ntl/: gentle.man, /rst/: karst.

²⁹ 出口 (1997: 18-19) のスペイン語圏 4 大都市の口語スペイン語コーパスの調査によれば、CV は 51.2%, CVC は 18.2% であった。表-11 b は書き言葉であるが、口語では CV の割合が少し増して CVC が減少している。

に C を加えて導く (16b)。いずれにしても 2 回の編集操作が必要である。

- (16) a. $CV_ +C \rightarrow \underline{CVC} -C \rightarrow VC$
b. $\underline{CV} -C \rightarrow V_ +C \rightarrow VC$

このように A という形態から B という形態に至る編集の操作数は A と B の間の「編集距離」(Edit distance) と呼ばれ³⁰、計量方言学 (dialectometría) で語形の類似度を計算するときに用いられる (Heeringa 2004, Nerbonne et al. 2010)。上の表の ED 列は最初のパターン (CV) とそれぞれのパターンとの編集距離を示している³¹。この表を見ると、最初のパターン (CV) から下に進むと次第に編集距離が長くなっていることが多いようだ。一部の例外があるが、そのときでも本来の距離と実際の距離の差は 1 に限られている。よって、それぞれの音節パタンの出現頻度は CV との編集距離 (E) とおおよそ反比例していることになる。つまり、CV からの編集距離が大きいほど出現頻度が少なく、有標性が高くなることが確認できる。

編集距離が同じ場合は、CC, CCC のように C が連続すると頻度が低くなる (17a,b)。

- (17) a. $\underline{C} V \underline{C} > \underline{CC} V > V \underline{CC}$
b. $\underline{CC} V C > \underline{CCC} V > V \underline{CCC}$

その場合、C の連続が V の前にあるパターンが V の後にあるパターンよりも頻度が高い (18)。

- (18) $\underline{CC} V \underline{CCC} > \underline{CCC} V \underline{CC}$

このように音節パターンは多要素になるほど (CV > CVC), そして複雑になるほど (CVC > CCV), 頻度が低くなり、有標性が増加する。その例外として CV > V があるが、これは編集距離という概念を導入すれば、V は CV に「削除」という編集操作が加わった結果であるので、その複雑性は増している、と考えれば仮説-I (5h) の反例とはならない³²。

ラテン語とスペイン語のペア 300 組の音節パターンを比較した Guffey (2002: 33-34) と現代スペイン語の書き言葉を調査した Navarro (1966: 47), そして出口 (1997: 18-19) の口語スペイン語コーパスの資料を整理した寺

³⁰ 付加・削除のほかに置換という操作が必要なこともある。ここでは必要でない。

³¹ R 関数 adist を使用した。

³² たとえば、中世スペイン語で頻繁に起きた語末母音の脱落などは、語形は短縮されているにも拘わらず、有標と見なされる (例: *parte* > *part*, *Fernando* > *Fernand*)。

崎 (2017: 72)³³ によると頻度の順位は次のようになる³⁴。

- (19) a. Guffey. ラテン語: CV > CVC > CCV > VC > V > CCVC > CVCC > VCC > CCVCC > CCCV > CCCVC.
b. Guffey. スペイン語: CV > CVC > CCV > VC > VC > V > CCVC > CVCC.
c. Navarro. 書き言葉コーパス: CV > CVC > V > CCV > VC > CCVC > VCC > CVCC > CCVCC.
d. 出口・寺崎: CV > CVC > V > VC > CCV > CCVC.

このように、表-11は(19)と類似しているので、両方の音節パターンは偶然ではなく、何らかの原則に基づいていると考えられる。

多数の言語を比較研究した Gordon (2016: 84-87) は音節パターンの有標性を次のようにまとめている。

- (20) a. Less marked » More marked: CV » CCV » CCCV
b. Less marked » More marked: CV » V

上の式(20a,b)は表-11と一致するので、スペイン語の辞書(D)とテキスト(E, O)に見られる音節パターンは一般的である、と言える。そして、表-11, (19), (20)の一致は音節パターンが有標性の序列(仮説-I)に基づいていることを示している。

2.2.2. 音節の中の音素

言語一般に音素が音節の中で従う調音時の開口度 (Saussure 1916/1971: 100-105) や聴覚的な聞こえ度 (Jespersen 1913: 190-193) の序列が知られている。Harris (1983: 15) は「聞こえ度の尺度」(sonority scale) に従ってスペイン語の音節に見られる次のような序列を挙げた。

- (21) [O]bstluent - [N]asal - [L]iquid - [G]lide - [V]owel

³³ (15d)では寺崎のCSVとCCV, CSVとCCVCをそれぞれCCV, CCVCにまとめた。

³⁴ (14a, b)の両言語の音節パターンは類似しているが、詳しい数値を見ると、ラテン語のCV(55.0%), CVC(28.5%)よりもスペイン語のCV(57.0%), CVC(18.0%)となっているので、スペイン語のCV > CVCのほうが強調されている。

一方, Zec (2007: 178) は一般音韻論の見地から次のように聞こえ度を細分化している (G を含めない)³⁵。

- (22) O voiceless stops
 voiceless fricatives
 voiced stops
 voiced fricatives
 N nasals
 L laterals
 rhotics
 V high vowels
 mid vowels
 low vowels

(21), (22) に従えば, スペイン語の音節は基本的に次のような構成 (テンプレート) になる。

- (23) [C1:C - C2:L /l,r/ - C3:G /y,w/]cx - V - [C4:G /y,w/ - C5:C - C6:/s/]cy

上のテンプレートの C1 と C5 の位置にさまざまな子音が入るが, C6 は /s/ に限られる。このテンプレートは V を中心としてかなり対称的な構成になっているが, 語頭の C1:C + C2:L /l, r/ は /p, t, k, b, d, g, f/ + /l, r/ という連続であり (tl, dl は除く), C5+C6 は /b, d, k, m, n, r/ + /s/ なので, 対称性は部分的に崩れている。また, /bs/ (subs.tan.cia), /ks/ (sex.to) は Zec の聞こえ度の序列 (22) に違反する。しかし, 丁寧な発音では /bs/, /ks/ は保たれるが, 普通の発話ではどちらも /s/ になるので聞こえ度の序列 (21, 22) に従っている。

寺崎 (2017: 69-70) によれば, 音節尾部は語末と語中で異なる。語末では /d, n, r, l, k, θ, s/: *piedad, afán, acabar, leal, cruz, lejos. 語中では /n, r, l, θ, s/: *cuento, orden, alba, bizcocho, pista. 学識語では /p, b, t, d, k, g, f/ がある : *apto, objeto, énrico, admirar, activo, digno, nafta.***

表-22, 23, 24 (→【付録】) は C1, C2, C3, V, C4, C5, C6 の位置に現れた音素とその出現個数を示している。V (母音) 以外の列のアステリスク (*) はゼロ (出現しないこと) を示す。

次は C1 に現れる子音 (24a,b,c) と阻害音 ([O]bstruent: 閉鎖音・破擦

³⁵ Harris (1983: 15) の左右の順番と Zec (2007: 178) の上下の順番はそれぞれここで示した順番と逆になっているが, これは私が語頭→母音という連続を考えたためである。

音・摩擦音) と共鳴音 ([S]onorant: 鼻音・流音・滑脱音) の区別 (24a',b',c') を頻度順で示す。

- (24) a. C1 (D): t > k > d > b > m > r > p > n > l > θ > s > g > ř > f > x > č > ʎ > ñ > y > w.
 b. C1 (E): d > t > k > l > s > p > m > b > n > r > θ > g > f > x > ř > ʎ > y > ñ > č > w.
 c. C1 (O): k > t > d > s > n > m > p > b > l > r > θ > g > x > y > f > č > ř > ʎ > ñ > w.
 a'. C1 (D): O > O > O > O > S > S > O > S > S > O > O > O > S > O > O > O > S > S > S > S.
 b'. C1 (E): O > O > O > S > O > O > S > O > S > S > O > O > O > O > S > S > S > S > O > S.
 c'. C1 (O): O > O > O > O > S > S > O > O > S > S > O > O > O > S > S > O > S > S > S > S.

(24a,b,c) の順番は、仮説-I の無標»有標の序列 (5) にほとんど一致する³⁶。そして、(20a',b',c') の阻害音 (O) と共鳴音 (S) の区別を見ると、序列の左側 (高頻度) で阻害音 (O) が多く、右側 (低頻度) で共鳴音 (S) が多くなっている。

さらに、C2, C3, V, C4 のそれぞれの位置で無標»有標の序列 (5) に従った次の頻度順が確認できる。

- (25) C2: (D, E, O) r > l, 反例なし。
 C3: (D, E, O) y > w, 反例なし。
 V: (D) a > o > e > i > u / (E, O): e > a > o > i > u,
 反例: o > e (D), e > a (E, O)
 C4: (D, E, O) y > w, 反例なし。

次は C5 に現れる子音と阻害音 ([O]bstruent) と共鳴音 ([S]onorant) の区別を頻度降順で示す。

- (26) a. C5 (D): n > r > s > l > k > d > y > w > θ > p > b > g > t > m > f > x > č.
 b. C5 (E): s > n > r > l > k > d > y > θ > w > b > g > p > t > m > x.
 c. C5 (O): s > n > r > l > y > d > m > k > θ > w > t > p > b > g > f.
 a'. C5 (D): S > S > O > S > O > O > S > S > O > O > O > O > O > S > O > O > O.
 b'. C5 (E): O > S > S > S > O > O > S > O > S > O > O > O > O > S > O.
 c'. C5 (O): O > S > S > S > S > O > S > O > O > S > O > O > O > O > O.

C6: s は唯一の子音を示し特殊である (Harris 1983: 15)。

- (27) C6: [DLE, JCh] s

³⁶ 例外は *b > p (D), *d > t (D), *k > t (O), *l > r (O), *m > b (E, O), *θ > s (D) である。

C5 の位置 (26) では、無標»有標の序列 (仮説-I) に従わない多くの例外がある。たとえば、*n > r, *n > l, *n > d, *r > d, *l > d, ... などである。ここで (26a', b', c') の内容を見ると、左側の高頻度の領域で多くの共鳴音 (S) が見つかると、右側の低頻度の領域で多くの阻害音 (O) が見つかる (/s/ は例外)。このことは、母音の後の位置では、共鳴音が普通であり、阻害音が特殊であることを示している。これは先に見た母音の前の位置 (C1) とはちょうど逆になる (24a', b', c')。そこで、先述の Colina (2014: 86) に従って (12), 母音の後の位置では阻害音 (無標) » 共鳴音 (有標) という序列を逆転させ、共鳴音 (無標) » 阻害音 (有標) としなければならない。そうすると、先に見た調音様式の例外はすべて仮説-I (5d) の反例にならなくなる³⁷。

Saussure (1916/1972: 79-82) は音節頭位における外破音の連鎖は小さな開口度の音から大きな開口度の音へ移行し、逆に音節尾位における内破音の連鎖は大きな開口度の音から小さな開口度の音へ移行するという対称的な連続の説明をしている。よって [O]bstruent, [N]asal, [L]iquid, [G]lide の分類を使うならばスペイン語の音節は下の図式 (28a) のような左右対称の関係になるはずである (Hooper 1976: 199; 出口 1997: 32; Gordon 2016: 99)。

- (28) a. O - N - L - G - <V> - G - L - N - O
 b. [[[O (L)], N] - G]_{CX} - <V> - [G - C (s)]_{CY}

しかし、実際の音節は上のような連続にはならず (表-22, 23, 24: →【付録】), むしろ (28b) のような非対称の図式になる。そして、先述のように、C1 (24) と C5 (26) のそれぞれの要素を比べると C1 での高頻度音素は阻害音 (O: t, k, d, b) であるが、C5 の高頻度音素は共鳴音である (S: n, r, l)。この C1: (24) と C5: (26) の不一致と (28b) の非対称性は、共に有標性の観点から説明できる。つまり、音節の核である母音の前の位置の子音は一般に阻害音 (O) が好まれ、逆に母音の後の位置の子音は一般に共鳴音 (S) が好まれている (/s/ は例外)。よって、母音の前の位置では典型的な子音である阻害音が選ばれ、母音の後の位置ではむしろ母音に近い子音である共鳴音が選ばれる。このことは、音節の構造が「開口度：開く→音節核→閉じる」「呼気圧：増加→音節核→減少」「筋肉の緊張：増加→音節核→減少」(Saussure 1916/1972: 79-82, Granda Gutiérrez 1966: 55; Hala 1973: 32) という図式に従っているとすれば納得できる。そこで、先述のように、核

³⁷ 反例になるのは、*k > t, *p > t, *d > t などの調音点と無声・有声の有標性だけである。

母音の後の位置では無標:阻害音»有標:共鳴音とし、核母音の後の位置では無標:共鳴音»有標:阻害音とする。

ここで、(29a=23) を阻害音 (O) と共鳴音 (S) を使って書き直すと (29b) になり、さらに (29b) の阻害音 (O) と共鳴音 (S) の流れ (連続) をまとめると (29c) になる。

- (29) a. [C1:C - C2:L /l,r/ - C3:G /y,w/]_{Cx} - V - [C4:G /y,w/ - C5:C - C6:/s/]_{Cy}
 b. [C1: (O - S) - C2: (S) - C3: (S)]_{Cx} - V - [C4: (S) - C5: (S - O) - C6: (O)]_{Cy}
 c. (O) - (S) - <V> - (S) - (O)

このことは、音節のそれぞれの位置 (O, S, V, S, O) の範列 (paradma) に出現する音素の出現頻度順 (24, 25, 26) が、音節構造の連辞 (sintagma) の構成 (29c) に近似することを示している (図-8)。

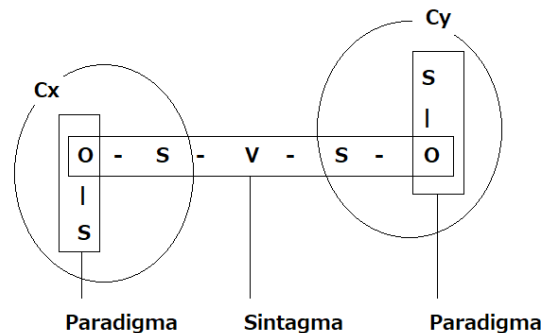


図-8. 音節の範列と連辞

これは偶然では起こらないことなので、このような縦列の範列的構成と横行の連辞的構成を繋ぐ同じ原則として、音素の使用頻度によって定位される有標性が考えられる。母音 (V) は最小の子音性を有するので、子音から見れば最も有標性が高い。母音に続くのが共鳴音 (S) であり、これが出現するときは母音に近い位置になる。性質的に母音から最も離れた阻害音 (O) は、音節の中で母音から最も離れた位置に現れる。よって、阻害音 (O) と共鳴音 (S) と母音 (V) の連辞的關係は (29c: O-S-V-S-O) 以外になく、たとえば、*S-O-V-S-O や *O-S-V-O-S のような音節は成立しない。そして、両端にある C1 と C5 の位置の音素の相対的頻度順位が C1: O > S, C5: S > O になる理由はそれぞれの有標性の序列が C1: O » S, C5: S » O になるためである。

2.2.3. 音素的音節

全体の音素的音節は辞書 (D) で 2353 種, 書き言葉資料 (E) で 1026 種, 話し言葉資料 (O) で 1096 種あるが, 次の表はそれぞれの上位 20 位を示す。

D	音節	頻度	有標度	E	音節	頻度	有標度	O	音節	頻度	有標度
1	a	10 930	0.000	1	de	44 790	0.000	1	a	14 743	0.000
2	te	8 441	0.228	2	a	38 565	0.139	2	ke	10 478	0.289
3	ta	7 697	0.296	3	la	22 455	0.499	3	no	8 224	0.442
4	ko	6 850	0.373	4	ke	20 745	0.537	4	de	7 382	0.499
5	ka	6 784	0.379	5	en	16 732	0.626	5	e	7 013	0.524
6	to	6 112	0.441	6	es	15 199	0.661	6	si	6 864	0.534
7	ti	5 509	0.496	7	el	14 617	0.674	7	es	6 148	0.583
8	do	5 257	0.519	8	se	14 269	0.681	8	te	5 694	0.614
9	ra	4 966	0.546	9	do	14 252	0.682	9	do	5 204	0.647
10	na	4 382	0.599	10	te	13 145	0.707	10	o	4 697	0.681
11	de	4 375	0.600	11	no	12 501	0.721	11	en	4 442	0.699
12	ro	4 217	0.614	12	ta	11 787	0.737	12	ko	4 389	0.702
13	ma	3 970	0.637	13	to	11 785	0.737	13	la	4 247	0.712
14	la	3 733	0.658	14	ra	11 579	0.741	14	ta	3 909	0.735
15	no	3 626	0.668	15	kon	9 794	0.781	15	se	3 905	0.735
16	mo	3 586	0.672	16	ko	9 576	0.786	16	ra	3 613	0.755
17	li	3 545	0.676	17	ka	9 370	0.791	17	na	3 528	0.761
18	ba	3 483	0.681	18	e	9 356	0.791	18	ro	3 419	0.768
19	da	3 465	0.683	19	na	9 351	0.791	19	to	3 294	0.777
20	pa	3 249	0.703	20	o	8 710	0.806	20	me	3 107	0.789

表-12. 音素的音節の頻度. 上位 20 位
(D: 辞書, E: 書き言葉, O: 話し言葉)

次は Moreno et al. (2006) による口語と書き言葉コーパスの音素的音節の頻度順である (降順)。

- (30) a. Moreno. Oral: a > ke > de > es > i > no > te > el > la > do.
b. Moreno. Escrito: de > a > la > ta > ke > en > do > te > na > *ma.

(30) にある音節のほとんどが表-12 に含まれているので, 高頻度の音素

的音節の頻度順はほぼ安定していると言える³⁸。

上位の音節はどれも開音節であり，ほとんどが CV 型であるが，テキスト (E, O) では en, es, el, kon という閉音節も見られる。これらの音節を構成する音素は無標のものばかりであり，頻度順は次のように無標»有標の序列 (仮説-I) に従っている。例外は少数である：*ba > pa (D); *de > te (E), *do > to (E), *la > ra (D, E, O)。

(31) a. D: a > [e, i, o, u], te > [de, pe, ke, če, ti], ta > [da, pa, ka, ča, ti, tu], ko > [go, čo, ki, ku], ka > [ga, ča, ki, ku], to > [do, po, ko, čo, so, θo, tu], ti > [di, pi, ki, či], do > [bo, go, du], ra > [la, řa, re, ri, ro, ru], na > [ma, ña, ni, nu], de > [be, ge, di], ro > [řo, lo, ru], ma > [ña, mi, mu], la > [ła, li, lu], no > [mo, ño, nu], mo > [ño, mu], li > łai, ba > [bi, bu], da > [ba, ga, di, du], pa > [bi, pi, pu].

b. E: de > [be, ge, ne, re, le, di], a > [e, i, o, u], la > [ła, li, lu], ke > [ge, če, ki], en > [in, em, eñ], es > [is, ef, ex], el > [il, eł], se > si, do > [bo, go, du], te > [pe, ke, ti], no > [mo, ño, nu], ta > [da, ti, tu], to > [po, ko, čo, so, θo, tu], ra > [řa, re, ri, ro, ru], kon > [gon, čon, kun, kom, koñ], ko > [go, ku], ka > [ga, ki, ku], e > i, na > [ma, ña, ni, nu], o > u.

c. O: a > [e, i, o, u], ke > [ge, če, ki], no > [mo, ño, nu], de > [be, ge, ne, re, le, di], e > i, si > [fi, θi, xi], es > [is, e, eθ, ex], te > [de, pe, ke, če, ti], do > [bo, go, du], o > u, en > [in, em, eñ], ko > [go, ku], la > [ła, li, lu], ta > [da, ti, tu], se > [fe, θe, xe, si], ra > [řa, re, ri, ro, ru], na > [ma, ña, ni, nu], ro > [řo, ru], to > [po, ko, čo, so, θo, tu], me > mi.

次は辞書(D), 書き言葉 (E), 話し言葉 (O) の資料中の音節のリスト (頻度降順) の先頭部分と末尾部分の抜粋である。

(32) a. (D): a te ta ko ka ... θyun ñarp ñaθ ñiθ ñur ñuθ ñya ñyer.

b. (E): de a la ke en ... řid trays treys xays yays ñes ñis.

c. (O): a ke no de e ... xyan xyas ying yon θays θid θyes.

このリストの末尾に近い音節は，どれも有標の音素が多く，複雑な構造を持ち，多数の音素が使われているので有標性が高い。したがって，このような膨大な音節のリストでも，高頻度音節が無標であり，低頻度音節が有標である，と考えられる。

このように，全体的傾向として，頻度が減少 (有標性が上昇) するに従って音節の大きさ (音素数) が増加しているように見える。そこで，頻度

³⁸ (30) の *ma は一致しない音節である

の減少（有標性の上昇）と音素数の関係を調べるために、出現頻度降順で並べた音節の音素数を計算し、順位（Ord）を10個のグループに分割して、両者をクロス集計した（表-13a,b）。

D	Ord	1	2	3	4	5	6	E	Ord	1	2	3	4	5	6
1	1	5	101	124	4	0	0	1	1	5	65	29	2	0	0
2	236	0	18	185	31	1	0	2	103	0	24	66	12	0	0
3	471	0	10	169	56	0	0	3	205	0	11	76	15	0	0
4	706	0	10	121	99	5	0	4	307	0	10	71	21	0	0
5	941	0	6	115	107	7	0	5	409	0	5	63	32	2	0
6	1176	0	5	97	117	15	1	6	511	0	7	64	29	2	0
7	1411	0	3	96	121	15	0	7	613	0	3	54	43	2	0
8	1646	0	8	83	124	20	0	8	715	0	3	48	44	7	0
9	1881	0	3	66	136	29	1	9	817	0	4	50	40	8	0
10	2116	0	1	81	128	24	1	10	919	0	2	33	52	14	1

表-13a,b. 音節の頻度順位 x 音素数: a. D.辞書 / b. E.書き言葉, Ord: 順位

O	Ord	1	2	3	4	5
1	11	5	61	37	5	0
2	110	0	27	74	8	0
3	219	0	16	74	19	0
4	328	0	8	69	31	1
5	437	0	10	60	37	1
6	546	0	7	67	35	0
7	655	0	3	57	46	3
8	764	0	4	64	38	3
9	873	0	7	43	50	9
10	982	0	4	53	45	7

表-14. 音節の頻度順位 x 音素数: O.話し言葉, Ord: 順位

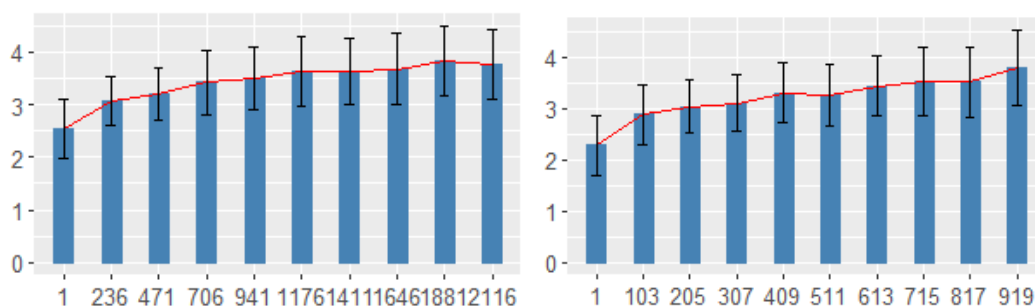


図-9a,b. 音節の頻度順位 x 音節数 (平均と標準偏差)

a. 辞書 (D) / b. 書き言葉テキスト (E)

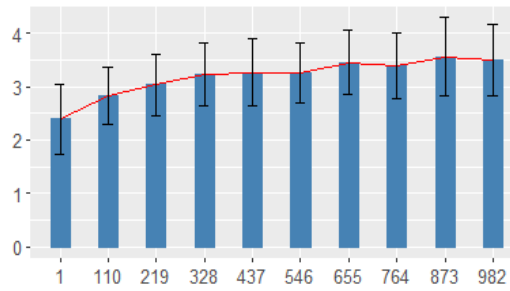


図-10. 音節の頻度順位 x 音節数 (平均と標準偏差)
話し言葉テキスト(O)

上の表-13, 14 と図-9, 10 を見ると, たしかに頻度が減少 (有標性が上昇) するに従って音節の大きさ (音素数) が増加していることが確認できる。よって音素の有標性の序列 (仮説-I. 5h) 「音素数: 少数»多数」をほぼ支持できる。

2.3. 語

音素的語が有する音素の数と音節の数に規則性や傾向があるのだろうか。はじめに, 3 種の資料の中の全出現語について音素数を数えると, 次のような頻度分布が得られた。

音素数: D	頻度	有標度	音素数: E	頻度	有標度	音素数: O	頻度	有標度
1	15	0.999	1	9	0.997	1	31	0.979
2	113	0.992	2	48	0.984	2	93	0.937
3	555	0.962	3	168	0.943	3	236	0.841
4	2 650	0.817	4	784	0.732	4	686	0.538
5	5 795	0.599	5	1 679	0.426	5	1111	0.252
6	10 459	0.276	6	2 339	0.201	6	1352	0.090
7	12 802	0.114	7	2 926	0.000	7	1486	0.000
8	14 443	0.000	8	2 890	0.012	8	1272	0.144
9	12 920	0.105	9	2 252	0.230	9	1014	0.318
10	9 882	0.316	10	1 674	0.428	10	646	0.565
11	6 925	0.521	11	993	0.661	11	384	0.742
12	4 436	0.693	12	453	0.845	12	203	0.863
13	2 927	0.797	13	226	0.923	13	118	0.921
14	1 706	0.882	14	106	0.964	14	61	0.959
15	993	0.931	15	47	0.984	15	32	0.978
16	480	0.967	16	12	0.996	16	9	0.994
17	226	0.984	17	1	1.000	17	8	0.995
18	117	0.992	20	1	1.000	18	6	0.996
19	42	0.997				19	2	0.999
20	20	0.999				20	1	0.999
21	6	1.000				21	2	0.999
22	2	1.000						
23	1	1.000						

表-15a,b,c. 語の音素数の分布 (a.D.辞書 : b.E.書き言葉 : c.O.話し言葉)

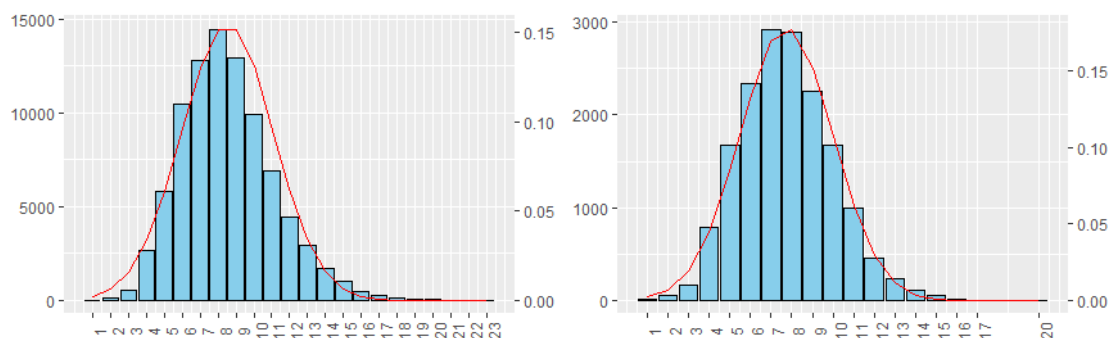


図-11-a,b. 語の音素数の分布 (a.D. 辞書 : b.E.書き言葉)

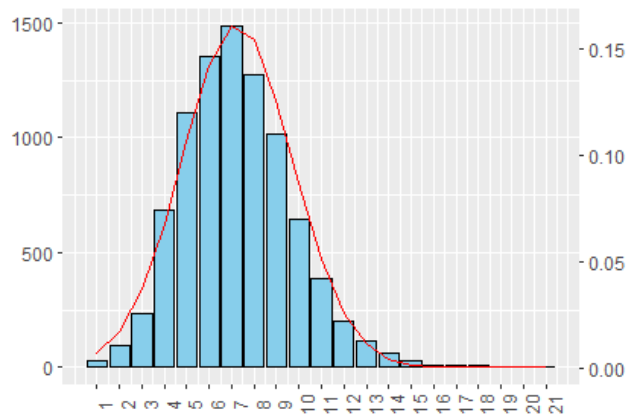


図-12. 語の音素数の分布 (O.話し言葉)

次は語の音節数の分布である。

音節数.D	頻度	有標度	音節数.E	頻度	有標度	音節数.O	頻度	有標度
1	685	0.977	1	163	0.976	1	290	0.913
2	10 681	0.636	2	3 438	0.491	2	2419	0.273
3	27 825	0.052	3	6 755	0.000	3	3327	0.000
4	29 356	0.000	4	4 833	0.285	4	1954	0.413
5	13 761	0.531	5	1 240	0.816	5	589	0.823
6	4 145	0.859	6	161	0.976	6	105	0.968
7	923	0.969	7	13	0.998	7	28	0.992
8	118	0.996	8	4	0.999	8	6	0.998
9	14	1.000				9	1	1.000
10	7	1						

表-16a,b,c. 語の音節数の分布 (a.D.辞書 : b.E.書き言葉 : c.O.話し言葉)

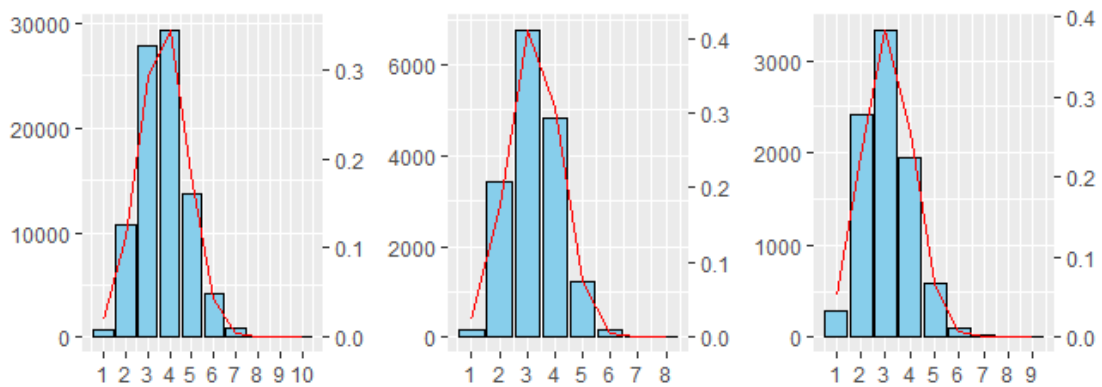


図-13-a,b,c. 語の音節数の分布 (a.D.辞書 : b.E.書き言葉 : c.O.話し言葉)

上の表と図を次のように観察する。

- (33) a. 音素数=8 (D), 7 (E, O) の語が最大の頻度を示し、それから両側（上側と下側）で離れるにつれて頻度が下がる。正規分布にかなり近似する単峰性の分布を示す。
- b. 音節数=4 (D), 3 (E, O) の語が最大の頻度を示し、それから両側（上側と下側）で離れるにつれて頻度が下がる。正規分布にやや近似する単峰性の分布を示す³⁹。

次のリストは表-16の語の音節数の頻度順 (34a,b,c) と Quilis (1981: 334) が示す話し言葉の語の音節数の頻度順 (34d) を比較したものである。

- (34) a. 音節数: D.辞書: 4, 3, 5, 3, 6, 7, 1, 8, 19, 10
- b. 音節数: E.書き言葉: 3, 4, 2, 5, 1, 6, 7, 8
- c. 音節数: O.話し言葉: 3, 2, 4, 5, 1, 6, 7, 8, 9
- d. 音節数: Quilis.話し言葉: 2, 1, 3, 4, 5, 6, 7, 8

本研究の主題である有標性と頻度との関係については、(仮説-I, 5h)「音素数：少数»多数」が関わる。図-11, 12, 13の音素数・音節数の頻度分布は単峰性を示し、単調減少を示していない。よって、語の音素数・音節数は有標性とは関係しない、と見られるかもしれないが、ここでも語の音素数=7~8, 音節数=3~4を無標とし、それから下降・上昇して離れるほど有標性が増す、と考えることもできる。むしろ、その方が有標性の概念が一般化するので好ましい。そのときは「下降有標性」(*marca descendente*)と「上昇有標性」(*marca ascendente*)を区別する。このように一般化した有標性を用いれば、有標性の増加は頻度の減少と相関する⁴⁰。

3. 応用⁴¹

次に、音素と音節構造の有標性理論を応用した歴史言語学の研究例を挙げる。近代初期に標準形 *así* と並んで *ansí* というバリエントが生起した理由を有標性を手掛かりとして探りたい。

³⁹ 竹林 (1996: 332) は『研究社：新スペイン語辞典』(1992)の基礎語彙 (*印のついた語, 4876語)の音節数の割合を調査し、3音節語が最大の頻度 (2032: 41.7%)を示すことを確認した。これを上の(33)の結果と比較すると、高頻度の基礎語彙では音節数が少なくなることがわかる。

⁴⁰ 単峰性の分布を示すデータであれば、すべて有標性の増加は頻度の減少と相関するが、複峰性の分布を示すデータでは相関しない。

⁴¹ 上田 (2022: 69-75) を抜粋・加筆した。

3.1. 年代

次のリストは Nieto Jiménez y Alvar Ezquerro (2007) と Real Academia Española (1726-1739) に収録された *así* / *ansí* の語形の歴史的分布を示す。

- (35) a. **así** : 1490, 1492, 1495, 1505, 1520, 1526, 1534, 1551, 1554, 1562, 1568, 1570, 1580, 1582, 1587, 1591, 1599, 1601, 1604, 1611, 1617, 1636, 1639, 1642, 1645, 1660, 1661, 1666, 1670, 1679, 1693, 1721, 1723, 1726.
- b. **ansí** : 1562, 1587, 1599, 1601, 1604, 1607, 1609, 1617, 1620, 1621, 1636, 1639, 1670, 1693, 1705, 1706, 1726.

このように辞書資料では *ansí* の初出は 1562 年である。Corominas y Pascual (1980: s.v. *así*) は中世・古典期に *ansí* が高頻度で用いられたと記述している。

3.2. 歴史

以下では *ansí* / *así* の歴史・地理・社会の状態を CODEA と COSER のコーパス資料を使って観察する⁴²。はじめに CODEA の資料を使って歴史的変化を見る。次の表の左側はそれぞれの絶対頻度を示す。右側は乗数を 10 万にしたときの正規化頻度である。グラフは正規化頻度で作成した⁴³。*ansí* が 1500 年代に上昇していることが確認できる。

絶対頻度	<i>ansí</i>	<i>así</i>	正規化頻度	<i>ansí</i>	<i>así</i>
1200	0	334	1200	0	249.8
1300	4	509	1300	1.9	242.6
1400	131	700	1400	47.5	253.9
1500	372	457	1500	128.3	157.6
1600	108	163	1600	60.6	91.4
1700	1	190	1700	0.7	139.5

表-17. *ansí* - *así*. 絶対頻度・正規化頻度. CODEA.

⁴² <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/codea.htm>
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/coser.htm>

⁴³ 年代: 総語数= 1200: 133,693, 1300: 209,804, 1400: 275,709, 1500: 289,897, 1600:178,337, 1700:136,183.

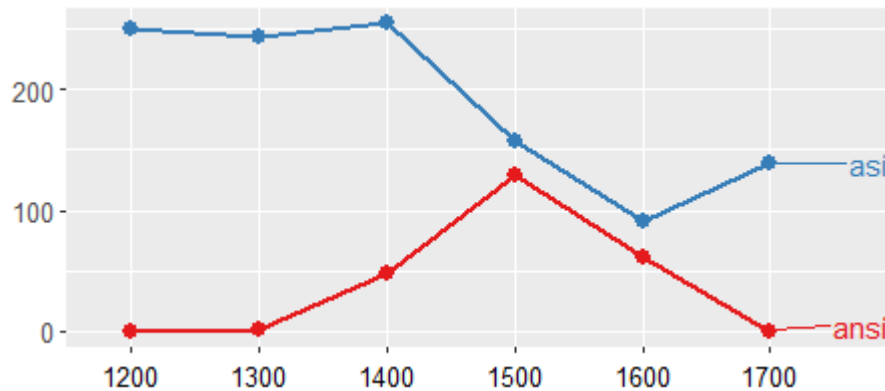


図-14. *ansi* - *asi*. 絶対頻度・正規化頻度・CODEA

現代のスペイン語の状況を見るために COSER コーパスを検索すると、半島中北部の Burgos に 1 例だけ見つかった。よって現代語で *ansi* は皆無に近いと言える。

3.3. 歴史・地理

次の表と地図は *ansi* の出現数の年代と地域のクロス集計の正規化頻度 (乗数: 10^5) を示している。*ansi* は 1500 年代の半島中央部で多く出現した。

FN. 10^5	1.PV	2.LE	3.CV	4.AR	5.CN	6.EX	7.AN
1300	0	3.8	3.5	0	2.7	0	0
1400	0	86.9	51.1	0	67.7	152	0
1500	18.4	186.1	117.2	49.1	143.2	366.4	79.1
1600	0	56.6	60.5	0	68.8	34.5	131.2
1700	0	0	0	0	2.6	0	0

表-18. *ansi* の出現数の年代と地域のクロス集計。正規化頻度 (乗数: 10^5)
(1.PV: País Vasco, 2.LE: León, 3.CV: Castilla la Vieja, 4.AR; Aragón,
5.CN: Castilla la Nueva, 6.EX: Extremadura, 7.AN: Andalucía.)

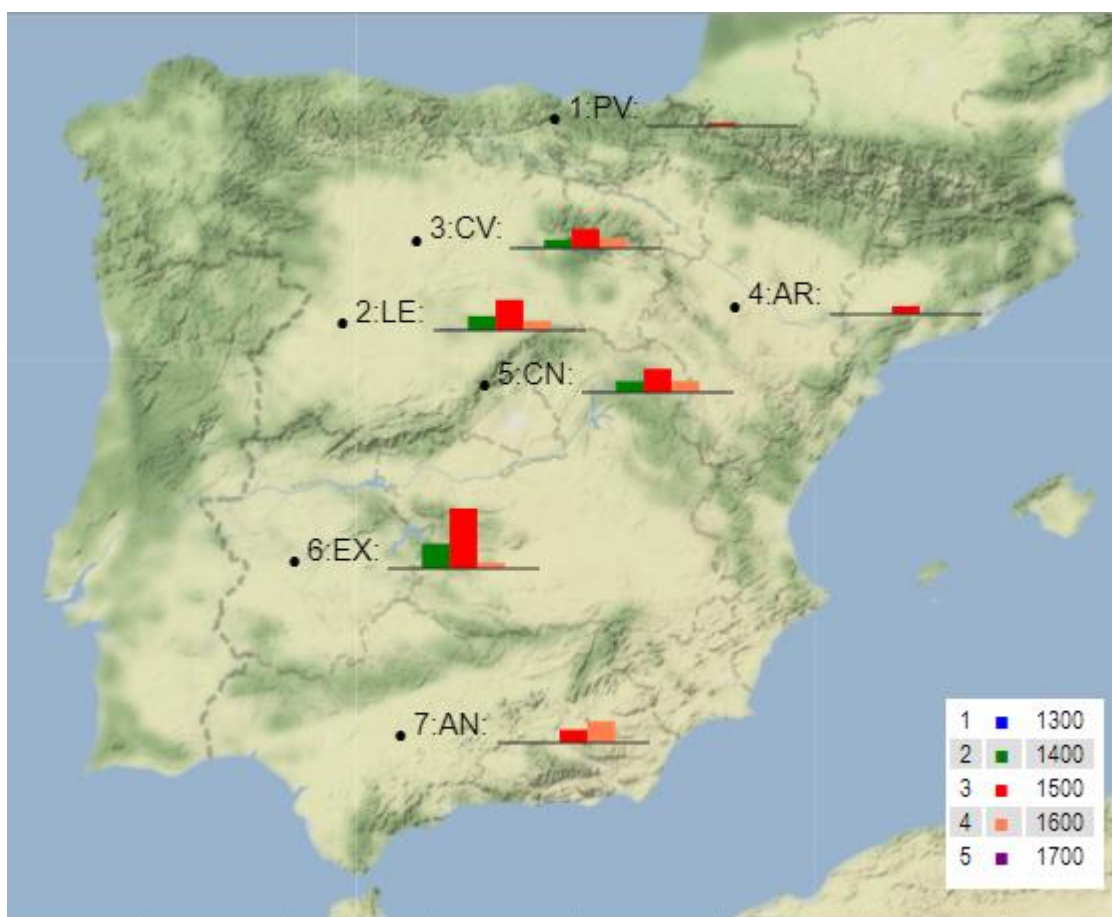


図-15. *ansí* の出現数の年代と地域. 正規化頻度⁴⁴
 (1.PV: País Vasco, 2.LE: León, 3.CV: Castilla la Vieja, 4.AR; Aragón,
 5.CN: Castilla la Nueva, 6.EX: Extremadura, 7.AN: Andalucía.)

3.4. 社会

次に王室文書・私文書を区別して両者間の社会・文体差を見る。

<i>ansí</i>	王室文書	私文書	<i>así</i>	王室文書	私文書
1200	0.0	0.0	1200	24.2	26.8
1300	0.2	0.3	1300	25.8	21.3
1400	2.3	5.2	1400	30.6	19.5
1500	9.2	11.5	1500	19.1	20.3
1600	2.4	5.0	1600	10.6	7.9
1700	0.0	0.1	1700	0.0	13.4

表-19. *ansí* - *así*. 王室文書・私文書. 正規化頻度. CODEA.

⁴⁴ <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/codea.htm>

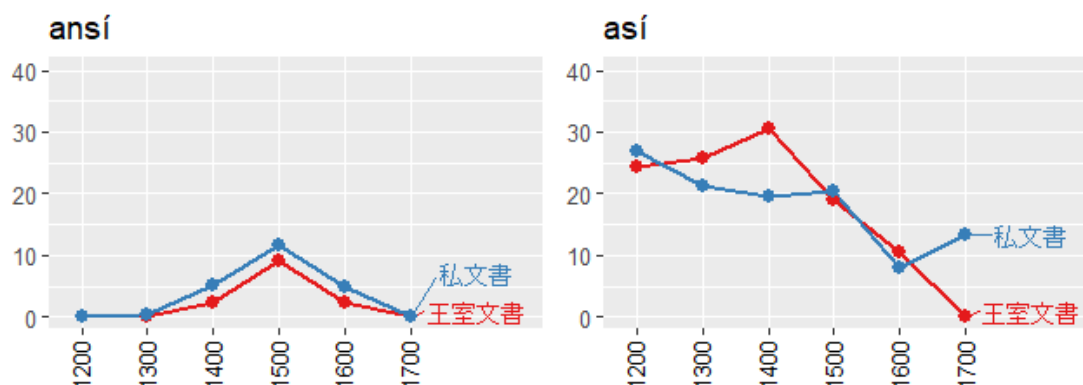


図-16. ansí - así. 王室文書・私文書. 正規化頻度. CODEA.

このように特定の年代・地域に現れる *ansí* は社会・文体の特徴はなく、王室文書と私文書の違いは検出されなかった。

3.5. 考察

así の語源は Rodríguez Molina (2015: 1053) によればラテン語形 AD SIC であり、一方、Corominas y Pascual (1980: s.v. *así*) によれば、*así* の語頭の *a-* は副詞に見られる単なる延長であった、という。筆者は、それに加えて SIC > [副詞] *sí* や接続詞 SI > *si* と語形が同じになるのでその同音衝突を避けた可能性があると考え。また、一般に語形が短いときは補強のために延長されることがあったことも原因の一つとして考えられる。この *así* は当初から全時代を通じて圧倒的多数であった。

一方、*ansí* については、筆者は無意味な *-n-* が音節末に生じた結果と考える。類例として *mucho* > *muncho*, *estonces* > *entonces*, *desde* > *dende* などがある⁴⁵。他にも LOCUSTA > *langosta*, MACULA > *mancha*, HIBERNU(S) > *invierno*, *maçana* > *manzana* などの変化に「理由不明の *-n-*」(Menéndez Pidal, 1968: 189), 「偶然の *-n-*」(Rodríguez Molina, 2015: 1055) が見られる⁴⁶。このような無意味の (偶然的な) *-n-* は *ansí* だけでなく *asín*, *asina*, *ansín*, *ansina* のように位置が変わったり繰り返されたりすることもある。音節末・語末の *-n-* は頻度が高いので、ほとんど無標の要素に近かったと思われる。また、音節末子音の情報量は小さいので音声変化が起こりやすい (上田 1975)。この *-n-* と前置詞 *en* の関係は不明なので、その影響を認める必

⁴⁵ *estonces* > *entonces*, *desde* > *dende* は 16 世紀の文人 Juan de Valdés (1535; ed. Lope Blanch 1969: 101) が挙げている例である。彼は *assí* > *ansí* も含めて、この *-n-* よりも *-s-* のほうがよい、と述べている。

⁴⁶ 中尾 (1996: 150) は [英] *messenger* (< *message*), *nightingale* (< ME. *nihtegale*) の例を載せている。

要はない。

Alvar y Pottier (1983: 339) は *ansí* が *asín, ansina, asina* などと共に俗語 (vulgar) であった, と述べている。逆に Del Barrio de la Rosa (2017: 254) は 1581-1620 の間 *ansí* は *así* よりも社会的に上位の語であった, という。また, 1500, 1600 年代の書簡や日記などの私文書を調べた Blas (2021) は *ansí* が低社会層で使われていた, と述べている。筆者は CODEA の資料によれば 1400, 1500, 1600 年代で私文書とともに王室文書でも使われていたので当時は両者に社会・文体差はなかったと考える (図-16)。社会・文体差が生まれたのは 1700 年代以降であり, その理由は *ansí* が低頻度で地域的に限定されていたため有標となり低く評価され, 広域・高頻度 (無標) の標準形 *así* の威信が高められたことによる (Rodríguez Molina 2015: 1051)。

よって, スペイン語史上 *así* > *ansí* > *así* という変化はなく, 常に *así* が優勢 (無標) であった。図式的に示せば両者の関係は次のようになる (A: *así*, B: *ansí*, A : B は A が優勢であったことを示す)。

A → A : B → A

次の表は 1500 年代の Castilla la Nueva における音節末子音の頻度と有標度を示している。

N	C.	Frec	GM	N	C.	Frec	GM	N	C.	Frec	GM
1	*	67 069	0.000	11	ns	80	0.999	21	nt	2	1.000
2	n	10 702	0.840	12	g	72	0.999	22	ms	1	1.000
3	s	9 893	0.852	13	m	55	0.999	23	ps	1	1.000
4	r	4 948	0.926	14	t	33	1.000	24	ts	1	1.000
5	l	3 357	0.950	15	f	31	1.000	25	v	1	1.000
6	d	1 052	0.984	16	b	19	1.000	26	x	1	1.000
7	θ	485	0.993	17	bs	9	1.000				
8	k	209	0.997	18	rs	6	1.000				
9	ks	163	0.998	19	rt	4	1.000				
10	p	97	0.999	20	ɫ	2	1.000				

表-20: 音節末子音の頻度 (Frec) と有標度 (GM)
(1500 年代の Castilla la Nueva, CODEA)

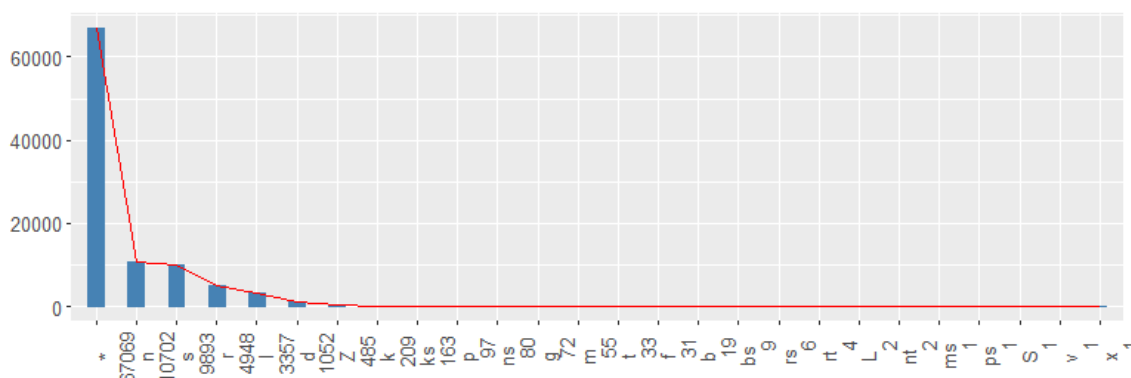


図-17. 音節末子音の頻度 (1500年代の Castilla la Nueva, CODEA)

上の表と図が示すように、音節末子音は無標の /ゼロ/, すなわち開音節に対比してすべて強い有標性を示していた。有標の子音の中で最も頻度が高かったのは /n/ (音節末子音の中では無標項) であったので挿入される可能性があった (他の子音は有標度が高いので生起しなかった)。しかし *ansí* は、子音の挿入自体が有標であり、その地理的分布が限定的であり、使用頻度が比較的低かったため、1500年代後半に移動した新首都 Madrid で標準形に成り得なかった。

序論で扱った「有標度を考慮した最適性理論」を適用すれば最適性のタブローは次のようになる。

	*ADD	*DEL	*CODA (GM)
a. a .sí			
b. an.sí	*		* (.840)
c. as.sí	*		* (.852)

表-21. /a.sí/, /an.sí/, /as.sí/

一般の最適性理論に従えば表-21 は認められない。an.sí と as.sí は表の最初にある *ADD (=DEP-IO, 追加の禁止) に違反しているからである。それを解決するには「制約の順番を変える」(*CODA→*ADD の順), という手段が採られる。筆者が提案する「有標度を考慮した最適性理論」では制約の順番を不動とする。そして、決定的な違反 (!) をカテゴリカルに表示せず、すべての違反の程度を連続的な有標度を使って表示する。とくに「忠実性の制約」(追加・削除を禁止) を重視して左側に位置させるが、これは絶対的な禁止ではなく、条件によって無視されることがある⁴⁷。ここでは「音節末子音の禁止」(*CODA) が関係す

⁴⁷ 「追加禁止」の制約の方が「削除禁止」の制約よりも優先する。これは言語一般に認められる「最小努力の法則」によって追加の方が削除よりも

る。表-20 によれば音節末の子音・子音結合には様々な有標度がある。その中で /n/ と /s/ の有標度が比較的低い (.840, .852)。そもそも最初の制約である「追加の禁止」に違反しているので、両者とも出現する可能性が小さいのだが、それでも有標度が一番低い /n/ が追加されることが稀にあった⁴⁸。/n/ に続く /s/ は音節末の位置に優先されるはずの共鳴音ではないため、非常に特殊な音素である。よって、それが自由に追加されることはなかった。

以上の議論では、頻度に基づく有標性仮説を歴史的変化の説明に応用した。当然、頻度に基づく有標性仮説を支持するためには、さらに多くの研究例を積み上げなければならない。

4. 結論

序論で仮説-I として掲げた音素・音節の有標性の序列 (5) と仮説-II の有標度の数式 (6) を以下に再掲する。

仮説-I: スペイン語音素の有標性の序列

- (5) a. 母音/半母音 : a » e » o » i » u » y » w:
 a > (e, o, i, u, y, w); e > (o, i, y); o > (u, w); i > (u, y); u > w
- b. 声帯振動 : 無声子音 » 有声子音: p > b; t > d; k > g.
- c. 調音位置 : 歯音/歯茎音 » 唇音/軟口蓋音 » 硬口蓋音
 t > (p, k) > č; d > (b, g); (θ, s) > (f, x); n > (m, ñ); m > ñ; l > ʎ.
- d. 調音様式 : 閉鎖音 » 破擦音 » 摩擦音 » 鼻音 » 流音:
 p > f; t > (θ, s); k > x; b > m; d > (n, r, l); n > (l, r, ř).
- e. 集約性 » 拡散性 : s > θ.
- f. 中断性 » 連続性 : r > l.
- g. 弛緩性 » 緊張性 : r > ř.
- h. 少数 » 多数 : t > tr; CV > CVC > CVCC など.
- i. 一般 » 特殊 : te.'nis.ta > 'Fró.mis.ta など.

稀であるためである。忠実性の制約には他に「変更の禁止」(*TRA) を加える。先述した「編集距離」(2.2.1) は追加・変更・削除を同じ回数で扱うが、この順番で離散的な比重 (3, 2, 1) や連続的な比重 (頻度) を加えるほうが精確に計算できる。連続的な比重を用いればさらに精密になる。

⁴⁸ /n/, /s/ の有標度が低いために、スペイン語のアクセント記号付与の規則では特別に扱われる。筆者は「語末が(母音, n, s)」という無標の強勢配置(語末から 2 番目の音節) の条件を「語末が母音(+n, s)」という条件にする。そこでは Vn, Vs が V と同じ価値を持つ (上田 2021)。

仮説-II: 有標度 (Grado de marcación: GM, F: 頻度, Max: 最大頻度):

$$(6) \quad GM = 1 - \frac{F}{Max}$$

本研究では、まったく異なる性質をもつ辞書の掲載語彙とテキストの使用語彙に現れた音素、音素的音節、音素的語という音韻的単位の中で、上に挙げた2つの仮説の有効性を検証した。音素の有標性と頻度の関係(仮説-I, 5a-g)の有効性はほとんどのケースで確認できたが、若干の例外も観察された。(仮説-I, 5h)については、音節パターンCVを無標とすれば、その他のパターン全体はわずかの例外を除いて、CVからの逸脱(編集距離)が大きくなるほど使用頻度が低くなることが検証された。また、膨大な音節の音素形式について、音節の長さ(音素数)は増加するにつれてその使用頻度が減少する傾向があることも確認できた。(仮説-I, 5i)については強勢の位置がラテン語の強勢配置規則に従わない例を見たが、実際のテキストに現れる蓋然性は極めて低いので、その有標性は完全に近い。

このように、(仮説-I, 5b, e, f, g)のように二項対立的に示された有標項の頻度は、ほとんどの場合低頻度であり、(仮説-I, 5a,c,d)のように離散的段階性があるときは、その段階性に依じて頻度が減少し、(仮説-I, 5h)のように有標性が連続的であるときは、有標性が漸増するに従って頻度が漸減する。そして、(仮説-I, 5i)のように有標性が極端に高い場合は、頻度も極端に低くゼロに近い。以上の観察から、スペイン語の音韻の有標性と頻度が一般にほぼ完全に逆相関するので、音韻の有標性を決定する際には、使用頻度が重要な基準となる。

有標性は言語体系の特徴であり、頻度は言語使用の結果であるので、基本的に両者の関係は「有標性が増加すれば頻度が減少する」ということになり、「要因：有標性→結果：低頻度」という方向が考えられる。しかし、音素的音節などのような低頻度の音韻的単位が、必然的にその有標性を示していることもある。この場合、xays, yays, ñes, ñisのような低頻度の音節が有標な単位として先験的に言語体系の中に存在するというよりも、言語使用の中に観察された低頻度性が有標性を与えている、と考えられる。よって、この場合は「要因：低頻度→結果：有標性」という方向性を示す。どちらの場合も「有標性≒低頻度」という本質的な関係が基本的に確認される。

一方、一部の例外もあることも確かである。実際の語彙目録(辞書)や言語使用(テキスト)の中で仮説-I(5)に掲げた有標性序列の例外が生じる理由は何であろうか。例外(有標項が無標項よりも頻度が高いケース)がある理由を筆者は次のように考察する。確かに無標項は調音の実現コス

トが低い（調音が簡単である）ので「最小努力の法則」(Principle of least effort, Zipf, 1949: 97-109) に従う。しかし、言語使用が「最小努力の法則」に従って無標項にばかり集中するならば、無標の音素や音節構造の使用が過剰になり、そのために音韻的単位の使命である弁別機能が阻害されてしまう。弁別機能を保持するためには、コストが高くても有標項も併せて使用しなければならない。そのうえ有標項は低頻度であるが故に情報量が豊富であるので軽視できない。そのときに有標項の頻度が上昇する、と考えられる。話し言葉のコミュニケーションにおいて話者は最小努力の法則に従って無標の音素や開音節型の音節パターンを使用しようとするが、聴者は最大弁別を求めて、有標の音素や閉音節型の音節パターンも期待する。実際の言語使用はこの両者のバランスの中で成り立っている。そのバランスが最大弁別（有標項）の方に強く傾いたときに有標性の例外が生じる、と考えられる。

以上の考察は音韻体系全体の一般的傾向を述べたものであるが、個々の音素や音節の頻度は使用される語の頻度に左右される。よって分析対象が異なれば、使用語彙も異なるので音韻的単位の頻度分布も異なることがある。たとえば Zipf (1936: 75) はスペイン語の /d/ は /t/ よりも頻度が大きいことを示しているが (/d/ > /t/, よって 仮説-I (5b) の反例になる), Zipf (1949: 102) では逆に /t/ > /d/ になっている（よって同仮説が支持される）。また、両者とも $p > b$ を示しているが、これは本研究で扱ったテキスト (E) の $b > p$ とは逆である。いずれにしても仮説-I (5) の反例は少数であるため、これらを考察から除外すればよいかもしれない。しかし、本研究では反例の頻度がたとえ少数であっても、頻度と有標度を考慮した上ですべて含める、という方策を採った。

スペイン語の音韻的単位（音素・音素的音節・音素的語）の有標性と頻度の関係を調べる時、仮説-II で示した「有標度」が役立つ。頻度を考慮しない有標性の理論において有標度は無用であるが、有標性の認識に頻度を考慮するならば、その数量的指標が必要である。本研究で提案した「有標度」を使って有標性の程度を計測し、有標性と頻度の相関だけでなく、その反例が起こる条件も特定することができた。

スペイン語は開音節型の言語である、と言われている (Malmberg 1965, Granda Gutiérrez 1966)。一方, Catalán (1971/1989: 78-79) は 12 世紀までのスペイン語はむしろ CVC 型であった、と言う。そして Guffey (2002) は O:阻害音（閉鎖音・摩擦音）と S:共鳴音（N:鼻音・L:流音・G:滑脱音）を区別し Hooper (1976) に従って O-N-L-G-V の左側を最適音節頭位とし、右側を最適音節尾位として、音節末に共鳴音を認める、という聞こえ度に基づく仮説を立て、ラテン語からスペイン語に至る様々な音韻変化を開音節

型優先の仮説と比較したところ，前者（聞こえ度に基づく仮説）の方が反例が少ない，ということを立証した。有標性と頻度の関係について，本研究のように現代語の資料だけでなく，歴史資料も含めて検証を続けなければならない。

参考文献

- Alarcos Llorach, Emilio. 1971. *Fonología española*. Madrid: Gredos.
- Albaigés Olivart, José María. 1989. *Diccionario de nombres de personas*. Universidad de Barcelona. Publicaciones.
- Alcina, J. y José Manuel Blecua. 1975. "Frecuencia de fonemas en español", en *Gramática española*. Barcelona: Ariel.
- Alvar, Manuel y Bernard Pottier. 1983. *Morfología histórica del español*. Madrid: Gredos.
- Andersen, Henning. 2001. "Markedness and the theory of linguistic change", in *Actualization. Current Issues in Linguistic Theory*, ed. by Henning Andersen. Amsterdam–Philadelphia: Benjamins, 21–57.
- Baquero, Julia M. 1999. *Fonología española. Hacia una teoría auténticamente generativa*. Bogotá: Universidad Nacional de Colombia.
- Blas Arroyo, José Luis (2021) "Apogeo y declive de así en los Siglos de Oro: nuevos datos desde la sociolingüística histórica", *Boletín de Filología*. vol. 56 no.1 versión On-line
<http://dx.doi.org/10.4067/S0718-93032021000100263>
- Boersma, Paul. 1998. *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives*. Doctoral dissertation, University of Amsterdam, The Hague.
<https://www.fon.hum.uva.nl/paul/papers/funphon.pdf>
- Catalán, Diego. 1971. "En torno a la estructura silábica del español de ayer y del español de mañana", en *Sprache und Geschichte. Festschrift für H. Meier*. München. Fink-Verlag, 78-220, recogido en *El español. Orígenes de su diversidad*. Madrid. Paraninfo, 77-104.
- Colina, Sonia. 2009. *Spanish Phonology. A Syllabic Perspective*. Washington, D. C.: Georgetown University Press.
- Colina, Sonia. 2012a. "Syllable Structure", Hualde, Ignacio José, Antxon Olarrea and Erin O'Rourke (eds.) *The Handbook of Hispanic Linguistics*. West Sussex: Wiley-Blackwell.

- Colina, Sonia. 2012b. "Phonotactic constraints on syllable structure", Colina, Sonia and Fernando Martínez Gil (eds.) *The Routledge Handbook of Spanish Phonology*, London: Routledge, 131-144.
- Colina, Sonia. 2014. "La sílaba en la teoría de la optimidad" en Hualde, José Ignacio. 2014, 86-89.
- Corominas, Joan y José A. Pascual. (1980) *Diccionario crítico etimológico castellano e hispánico*. 6 vols. Madrid: Gredos.
- Del Barrio de la Rosa, Florencio. (2017) "Piezas léxicas y variación morfosintáctica en la historia del español. Tres casos en el español de los Siglos de Oro (1581-1620)" en Del Barrio (ed.) *Palabras Vocabulario Léxico. La lexicología aplicada a la didáctica y a la diacronía*. Venezia: Edizioni Ca Foscari, 251-266.
- 出口厚実. 1997. 『スペイン語学入門』 東京 : 大学書林.
- D'Introno, Francisco, Enrique Del Teso, Rosemary Weston. 1995. *Fonética y fonología actual del español*. Madrid: Cátedra.
- Domínguez, Giuseppe. 2021. *Diccionario de la RAE en modo texto plano*.
<https://www.giuseppe.net/blog/archivo/2015/10/29/diccionario-de-la-rae-en-modo-texto-plano/>
- Fernández Sevilla, Julio. 2000. "Los fonemas implosivos en español", en J. Gil Fernández (ed.). *Panorama de la Fonología española actual*. pp. 207-234.
- Gordon, Matthew K. 2016. *Phonological Typology*. Oxford University Press.
- Granda Gutiérrez, Germán de. 1966. *La estructura silábica y su influencia en la evolución fonética del dominio ibeo-románico*. Madrid: Revista de Filología Española, anejo 81.
- Gil Fernández, Juana. 1990. *Los sonidos del lenguaje*. Madrid: Editorial Síntesis.
- González Rátiva, María Claudia y Mejía Escobar, Jorge Antonio. 2011. "Frecuencia de fonemas en dos corpus de español de uso en Colombia", *Actas de XVI Congreso Internacional de la ALFAL – Alcalá 2011*, 105-115.
- Granda Gutiérrez, Germán. 1966. *La estructura silábica y su influencia en la evolución fonética del dominio ibero-románico*. Revista de Filología Española, anejo 81.
- Guffey, Karen. 2002. *Spanish Syllable Structure*. Lanham: University Press of America.
- Guirao, Miguelina y María Amalia García Jurado. 1993. *Estudio estadístico del español*. Buenos Aires: Consejo Nacional de Investigaciones Científicas y Técnicas.

- Guirao, Miguelina y Ana Borzone de Manrique. 1972. "Fonemas, sílabas y palabras del español de Buenos Aires", *Filologia*, XVI: 135-165.
- Hayes, Bruce. 1998. "Gradient Well-Formedness in Optimality Theory"
<https://linguistics.ucla.edu/people/hayes/gradient.htm>
- Hara, Makoto. 1973. *Semivocales y neutralización*. Madrid: C.S.I.C.
- Harris, James W. 1969. *Spanish Phonology*. Massachusetts. The M.I.T. Press.
- _____. 1983. *Syllable Structure and Stress in Spanish*. The Massachusetts Institute of Technology.
- Heeringa, W. 2004. *Measuring Dialect Pronunciation Differences Using Levenshtein Distance*. PhD thesis, Rijksuniversiteit Groningen.
<https://pure.rug.nl/ws/portalfiles/portal/9800656/thesis.pdf>
- Hockett, Charles F. 1955. *A Manual of Phonology*. Baltimore: Waverly Press.
- Hooper, Joan B. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.
- Hualde, José Ignacio. 2014. *Los sonidos del español*. Cambridge University Press.
- Hyman, Larry M. 1975. *Phonology. Theory and Analysis*. New York: Holt, Rinehart and Winston.
- Jakobson, R. 1968. *Child Language, Aphasia, and Phonological Universals*. (trans. A. Keiler). The Hague: Mouton.
- Jakobson, Roman, C. Gunnar M. Font. Morris Halle. 1952. *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates*. Massachusetts Institute of Technology. 竹林滋・藤村靖訳. 1965. 『音声分析序説—弁別の特徴とその関連量』東京：研究社.
- Jakobson, Roman & Linda R. 1979. *The Sound Shape of Language*. Indiana University Press. 松本克己訳. 1986. 『言語音形論』東京：岩波書店.
- Jespersen, Otto. 1913. *Lehrbuch der Phonetik*. Leipzig: Teubner.
- Juillard, Alphonse & E. Chang-Rodriguez. 1964. *Frequency Dictionary of Spanish Words*. The Hague: Mouton.
- Kaisse, Ellen M. 2021. "Glides and high vowels in Spanish", Colina, Sonia and Fernando Martínez Gil (eds.) *The Routledge Handbook of Spanish Phonology*, London: Routledge, 145-161.
- 窪菌晴夫. 1998. 『音声学・音韻論』東京：くろしお出版.
- 窪菌晴夫・本間猛. 2002. 『音節とモーラ』東京：研究社.
- Lass, Roger. 1984. *Phonology. An Introduction to Basic Concepts*. London: Cambridge University Press.
- Llisterri, Joaquim y José B. Mariño. 1993. "Spanish adaptation of SAMPA and

- automatic phonetic transcription"
http://liceu.uab.cat/~joaquim/publicacions/SAMPA_Spanish_93.pdf
- Lope Blanch, Juan M. 1967. "La influencia del sustrato en la fonética del español de México", *Revista de Filología Española*, 50, 145-161.
- Lope Blanch, Juan M. (ed.) 1969. *Juan de Valdés: Diálogo de la lengua*. Madrid. Castalia.
- Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge University Press.
- Malmberg, Bertil. 1965. "La estructura silábica del español", en *Estudios de fonética hispánica*, Madrid, C.S.I.C, 3-28.
- Martinet, André. 1964. *Économie des changements phonétiques. Traité de phonologie diachronique*. Berne. Fuente Arranz, Alfredo de (trad. 1974) *Economía de los cambios fonéticos*. Madrid: Gredos.
- Martinet, André. 1965. *La linguistique synchronique. Etudes et recherches*. Presses Universitaires de France. 渡瀬嘉朗訳. 1977. 『共時言語学』東京：白水社.
- Martinet, André. (ed.) 1969. *La linguistique. Guide alphabétique*. Edition Denoël. 三宅徳嘉 (監訳). 1972. 『言語学事典. 現代言語学—基本概念 51 章』東京：大修館書店.
- Martínez, Inmaculada y Hiroto Ueda. 2021. *Inventario léxico de PRESEEA-Santander. Proyecto para el estudio sociolingüístico del español de España y América*.
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/chiri/inventario-santander.pdf>
- Martínez, Inmaculada y Hiroto Ueda. (en prensa) "Frecuencia de los fonemas en el corpus PRESEEA-Santander"
- Martínez Celdrán, Eugenio. 1994. *Fonética*. Barcelona: Teide.
- Menéndez Pidal, Ramón. 1968. *Manual de gramática histórica española*. 13 ed. Madrid: Espasa-Calpe.
- Moreno Sandoval1, Antonio, Doroteo Torre Toledano1, Natalia Curto y Raúl de la Torre1. 2006. "Inventario de frecuencias fonémicas y silábicas del castellano espontáneo y escrito", *IV Jornadas en Tecnología del Habla*, Zaragoza, 77-81.
<http://lorien.die.upm.es/~lapiz/rtth/JORNADAS/IV/4jth.pdf>.
- Mosonyi, Esteban Emilio. 1970. "Reseña de Bertil Malmberg, Estudios de fonética hispánica.", *Nueva Revista de Filología Hispánica*, 19, 425-430.
- Navarro Tomás, Tomás. 1966. *Estudios de fonología española*. New York: Las Américas Publishing Company.

- Navarro Tomás, Tomás. 1972. *Manuel de pronunciación española*. 18a ed. Madrid: C.S.I.C.
- Nerbonne, John, Jelena Prokić, Martijn Wieling & Charlotte Gooskens. 2010. "Some Further Dialectometric Steps", *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*, 41-56.
https://www.researchgate.net/publication/268337579_Some_further_dialectometric_steps/stats
- Nieto Ballester, Emilio. 1997. *Breve diccionario de topónimos españoles*. Madrid. Alianza Editorial.
- Nieto Jiménez, Lidio y Manuel Alvar Ezquerro. 2007. *Nuevo tesoro lexicográfico del español (S. XIV-1926)*. 11 vols. Madrid. Arco / Libros.
- Pensado, Carmen. 1985/2000. "Sobre la interpretación de lo inexistente: los tipos silábicos inexistentes en la fonología del español", *Folia Lingüística*, 19, 313-319, recogido en Gil Fernández, , Juana. (ed.). 2000. *Panorama de la fonología española actual*, Madrid: Arco / Libros, 475-483.
- Prince, Alan and Paul Smolensky. 1993/2002. *Optimality Theory. Constraint Interaction in Generative Grammar*. RuCCS-TR-2; CU-CS-696-93. ROA Version: August, 2002.
<http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF>
- Oller, D. K. and R. E. Eilers. 2008. "Similarity of Babbling in Spanish- and English-learning Babies", *Journal of Child Language*, 9, 565 - 577.
http://jth2006.unizar.es/finals/4jth_116.pdf
- Pérez, Hernán Emilio. 2003. "Frecuencia de fonemas", *e-rthabla*, 1.
http://lorien.die.upm.es/~lapiz/e-rthabla/numeros/N1/N1_A4.pdf
- Pineda, L. A., P. L. Villaseñor., J. Cuétara. H. Castellanos, I. López. 2004. "DIMEx100: A New Phonetic and Speech Corpus for Mexican Spanish", en C. Lemaître, C.A. Reyes and J.A. Gonzalez (eds.), *Iberamia*, 2004, LNAI 3315: 974–983.
<http://leibniz.iimas.unam.mx/~luis/DIMÉpublicaciones/papers/DIMEx100-LNAI3315.pdf>
- Quilis, Antonio. 1993. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Quilis, Antonio y M. A. Esgueva Martínez. 1980. "Frecuencia de fonemas en el español hablado", *Lingüística Española Actual*, 2. 1-25.
- R Core Team (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
<https://www.R-project.org/>.

- Real Academia Española. 1726-1739. *Diccionario de autoridades*. Madrid: Real Academia Española.
<https://apps2.rae.es/DA.html>
- Real Academia Española. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Rice, Keren. 2007. "Markedness in Phonology", in Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press.
- Rodríguez Molina, J. 2015. "El adverbio *así* en español medieval: variantes morfológicas", en García Martín, J. *Actas del IX Congreso Internacional de Historia de la Lengua Española (Cádiz, 2012)*, vol. 1. Madrid. Frankfurt am Main: Iberoamericana / Vervuert, 1049-1064.
- Rojo, Guillermo. 1991. "Frecuencia de fonemas en español actual", *Homenaje ó Profesor Constantino García*. Universidad de Santiago de Compostela: 451-467.
- Saussure, Ferdinand de. 1916 *Cours de linguistique générale*. Alonso, A. (trad.) 1971. Buenos Aires: Editorial Losada. Amado Alonso (trad.). 1971. *Curso de lingüística general*. Buenos Aires. Losada. 小林英夫訳 1972. 『一般言語学講義』 東京：岩波書店.
- Schane, Sanford A. 1973. *Generative Phonology*. Englewood Cliffs. Prentice-Hall. 桑原輝男・根間弘海訳 『生成音韻論』 東京. 研究社.
- 竹林滋. 1996. 『英語音声学』 東京：研究社.
- 田中伸一. 2009. 『日常言語に潜む音法則の世界』 東京：開拓社.
- 田中春美(編). 1988. 『現代言語学辞典』 東京：成美堂.
- 寺崎英樹. 2017. 『スペイン語文法シリーズ：発音・文字』 東京：大学書林.
- Torreblanca, Máximo. 1980. "La sílaba española y su evolución fonética", *Thesaurus* 35, 506-515.
- Trask, R. L. 2000. *The Dictionary of Historical and Comparative Linguistics*. Edinburgh University Press.
- Trubetzkoy, N. S. 1969. *Principles of Phonology*. University of California Press. 長嶋喜朗訳. 1980. 『音韻論の原理』 東京：岩波書店.
- 上田博人. 1975. 『スペイン語音節の音韻論的考察』 東京外国語大学卒業論文. <https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/onsei/onsetu.pdf>
- 上田博人. 2021. 「スペイン語の語尾と強勢：辞書・大規模コーパス・文体・社会地理・歴史地理資料の分析」 『地理言語学研究』 (日本地理言語学会), 1, pp. 51-105.
<https://geolinguistics.sakura.ne.jp/Journal/sig01.pdf>
- 上田博人. 2022. 「スペイン語高頻度語 «así», «mismo», «ahora» の地理・

社会的変異と歴史的変化の要因」『地理言語学研究』(日本地理言語学会), 2, 66-101.

<https://geolinguistics.sakura.ne.jp/Journal/sig02.pdf>

Zec, Draga. 2007. "The syllable", Lacy, Paul de (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press. 161-194.

Zipf, George K. 1936. *The Psycho-biography of Language. An Introduction to Dynamic Philology*. London: Routledge.

Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort. An Introduction to Human Ecology*. Mansfield: Addison-Wesley Press.

Zipf, George. K. y F. M., Rogers. 1939. "Phonemes and Variphones in Four Present-day Romance Languages and Classical Latin from the Viewpoint of Dynamic Philology", 15, 111-147.

【付録】

C1	頻度	GM	C2	頻度	GM	C3	頻度	GM	V	頻度	GM	C4	頻度	GM	C5	頻度	GM	C6	頻度	GM
t	41 298	0.000	*	305 178	0.000	*	305 255	0.000	a	109 103	0.000	*	323 855	0.000	*	232 822	0.000	*	323 086	0.000
*	33 715	0.184	r	14 063	0.954	y	15 079	0.951	o	76 024	0.303	y	206	0.999	n	33 370	0.857	s	1112	0.997
k	28 149	0.318	l	4 957	0.984	w	3 864	0.987	e	73 940	0.322	w	137	1.000	r	25 248	0.892			
d	27 175	0.342							i	49 733	0.544				s	15 507	0.933			
b	20 956	0.493							u	15 398	0.859				l	7 146	0.969			
m	20 471	0.504													k	3 024	0.987			
r	18 728	0.547													d	1 596	0.993			
p	18 199	0.559													y	1 251	0.995			
n	17 837	0.568													w	1 217	0.995			
l	17 546	0.575													θ	944	0.996			
θ	16 817	0.593													p	623	0.997			
s	16 770	0.594													b	390	0.998			
g	10 377	0.749													g	384	0.998			
ř	8 375	0.797													t	318	0.999			
f	8 215	0.801													m	254	0.999			
x	7 266	0.824													f	59	1.000			
č	4 756	0.885													x	27	1.000			
Ľ	3 870	0.906													č	17	1.000			
ñ	2 022	0.951																		
y	1 317	0.968																		
w	339	0.992																		

表-22. 音節内音素の出現頻度 (辞書: DLE), GM: 有標度

C1	頻度	GM	C2	頻度	GM	C3	頻度	GM	V	頻度	GM	C4	頻度	GM	C5	頻度	GM	C6	頻度	GM
*	145 330	0.000	*	782 008	0.000	*	771 129	0.000	e	282 076	0.000	*	820 791	0.000	*	535 589	0.000	*	819 602	0.000
d	100 724	0.307	r	32 481	0.958	y	35 955	0.953	a	235 526	0.165	y	861	0.999	s	97 277	0.818	s	2653	0.997
t	80 520	0.446	l	7 766	0.990	w	15 171	0.980	o	180 411	0.360	w	603	0.999	n	96 279	0.820			
k	68 627	0.528							i	79 890	0.717				r	39 481	0.926			
l	58 655	0.596							u	44 352	0.843				l	33 973	0.937			
s	56 979	0.608													k	6 177	0.988			
p	47 534	0.673													d	5 275	0.990			
m	45 863	0.684													y	3 427	0.994			
b	44 116	0.696													θ	1 747	0.997			
n	43 289	0.702													w	1 077	0.998			
r	33 075	0.772													b	760	0.999			
θ	29 127	0.800													g	503	0.999			
g	15 067	0.896													p	456	0.999			
f	11 828	0.919													t	117	1.000			
x	10 970	0.925													m	97	1.000			
ř	10 612	0.927													x	20	1.000			
Á	6 382	0.956																		
y	4 952	0.966																		
ñ	4 323	0.970																		
č	4 122	0.972																		
w	160	0.999																		

表:-23. 音節内音素の出現頻度 (書き言葉テキスト: JCh), GM: 有標度

C1	頻度	GM	C2	頻度	GM	C3	頻度	GM	V	頻度	GM	C4	頻度	GM	C5	頻度	GM	C6	頻度	GM
*	52 223	0.000	*	262 965	0.000	*	256 120	0.000	e	91 419	0.000	*	273 144	0.000	*	189 609	0.000	*	273 353	0.000
k	26 790	0.487	r	7 650	0.971	y	9 760	0.962	a	75 838	0.170	y	387	0.999	s	30 681	0.838	s	265	0.999
t	26 652	0.490	l	3 003	0.989	w	7 738	0.970	o	64 869	0.290	w	87	1.000	n	27 562	0.855			
d	22 849	0.562							i	27 879	0.695				r	11 993	0.937			
s	20 194	0.613							u	13 596	0.851				l	6 900	0.964			
n	18 036	0.655													y	2 208	0.988			
m	17 474	0.665													d	1 348	0.993			
p	16 681	0.681													m	1 126	0.994			
b	16 322	0.687													k	1 078	0.994			
l	13 465	0.742													θ	438	0.998			
r	10 696	0.795													w	231	0.999			
θ	7 717	0.852													t	168	0.999			
g	5 935	0.886													p	90	1.000			
x	4 187	0.920													b	77	1.000			
y	3 486	0.933													g	62	1.000			
f	2 766	0.947													f	33	1.000			
č	2 133	0.959																		
ř	2 024	0.961																		
ʎ	1 739	0.967																		
ñ	1 239	0.976																		
w	966	0.982																		

表-24. 音節内音素の出現頻度 (話し言葉テキスト: Santander), GM: 有標度