

Article

The Genesis of Spanish /θ/: A Revised Model

Ian Mackenzie

School of Modern Languages, Newcastle University, Newcastle upon Tyne NE1 7RU, UK;
ian.mackenzie@ncl.ac.uk

Abstract: This article proposes a revised model of the genesis of Castilian Spanish /θ/, based on (i) precise tracking across the Late Middle Ages of the orthographical *d* → *z* change in preconsonantal coda position and (ii) the potential for auditory indeterminacy between denti-alveolar variants of [s] and the non-sibilant [θ]. According to the findings, two non-sibilant phonemes, /θ/ and /ð/, are likely to have come into existence by the early 1500s, merger at the expense of /ð/ occurring shortly thereafter. This effectively inverts the normally assumed chronology, according to which devoicing preceded and indeed was implicated in the genesis of /θ/. The revised chronology weakens the teleological analysis of /θ/, which treats its genesis in terms of a functionally motivated widening of the articulatory distance between similar-sounding sibilants. Instead, the emergence of Castilian /θ/ is argued to be a natural reflex of the auditory permeability between the denti-alveolar type of [s] and the non-sibilant [θ], with analogous evolutions occurring outside the domain of Castilian Spanish. As part of this overall approach, the article assumes dissibilant (understood as the converse of assibilant) to be the fundamental process in the genesis of /θ/, rather than interdentalization.

Keywords: sound change; Castilian; Spanish; quantitative; S-curve; /θ/; sibilants; Old Spanish; *ceceo*; logistic regression



Citation: Mackenzie, Ian. 2022. The Genesis of Spanish /θ/: A Revised Model. *Languages* 7: 191. <https://doi.org/10.3390/languages7030191>

Academic Editors: Juana M. Licerias and Raquel Fernández Fuertes

Received: 13 March 2022

Accepted: 15 July 2022

Published: 22 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Castilian Spanish phoneme /θ/ represented by the *z* of *pozo* ‘well’ or the *c* of *vecino* ‘neighbor’ has often been considered to be something of an outlier within the general context of Romance phonological change. The ancestor of this sound is the affricated sibilant /ts/—geminated to /tts/ in certain contexts—which is assumed to have existed quite generally in early Romance, where it represented an assibilated continuation of the sequence [tj] that arose in various contexts in spoken Latin.¹ In the majority of cases, the modern reflex of early Romance /ts/ is a sibilant, still affricated in many varieties of Italo-Romance, but deaffricated and split into /s/ and /z/ in Gallo-Romance and much of Ibero-Romance. Considered in that light, the non-sibilant outcome of /θ/ found in Castilian and the adjacent languages Asturian and (standard) Galician looks markedly atypical. It should be borne in mind, however, that this type of result is not exclusive to northern Iberia. Thus, /θ/ occurs in conservative varieties of Venetian (Zamboni 1974), where its distribution resembles that of Castilian /θ/, and also in certain dialects of Franco-Provençal, although in the latter language it appears to be limited to contexts in which French has /ʃ/ (Hinzelin 2018). In addition, it should not be forgotten that in some varieties of Andalusian Spanish, the merged phoneme corresponding to Castilian /s/ and /θ/ has traditionally been capable of being articulated as [θ]—a phenomenon known as *ceceo*—although the relevant speakers typically alternate this pronunciation with more sibilant-like realizations (see Navarro Tomás et al. 1933; Villena Ponsoda et al. 1994–1995).

As regards the /θ/ of Castilian Spanish, its genesis has usually been approached from a teleological perspective (cf. Rost Bagudanch 2022), in the sense that the sixteenth-century fortition or devoicing of the Spanish coronal fricatives is usually analyzed as having created a Basque-style clustering of purely voiceless sibilants that was somehow in need of ‘resolution’, to use Widdison’s (1987, p. 67) term. Early exponents of this approach include

the functionalists [Martinet \(1951\)](#) and [Alarcos Llorach \(1988, 1991\)](#), both of whom analyzed the emergence of /θ/ as part of a structural reorganization aiming to re-establish auditory distance between phonemes that had come to be phonetically too similar to one another. Although not couched in explicitly functionalist terms, essentially the same explanation has been advanced by most modern scholars, as is described in more detail in Section 2. Anticipating the discussion in that section, it can be observed that the conventional model of the genesis of Castilian /θ/ involves two core assumptions, viz., (i) that the emergence of /θ/ postdates the devoicing of the coronal fricatives, and (ii) that the diachronic process that delivered /θ/ responded to a need to increase the articulatory distance between similar-sounding voiceless sibilants.

The present paper questions the foregoing approach, on the basis of quantitative data from the Late Middle Ages, relevant descriptive observations from commentators such as Antonio de Nebrija (1444–1522) and a reasoned critique of the teleological explanation for the sound change which gave rise to /θ/. The quantitative findings reported suggest that the standard chronology should be inverted, in the sense that the emergence of Castilian /θ/ seems likely to predate rather than postdate the devoicing of the coronal fricatives. As a corollary, it would follow that Spanish never had a clustering of purely voiceless sibilants that needed to be ‘resolved’—a scenario that leaves the teleological account with little in the way of independent motivation. A reanalysis-based explanation for the change, motivated by auditory indeterminacy, is therefore proposed as a more plausible alternative.

2. Previous Analyses

As was mentioned in the Introduction, the conventional or standard model for the genesis of /θ/ in Castilian Spanish has its roots in the functionalist thinking of authors such as André [Martinet \(1951\)](#) and Emilio [Alarcos Llorach \(1988, 1991\)](#). In Alarcos Llorach’s analysis, for example, the sixteenth-century devoicing of the Spanish coronal fricatives created a situation in which, in addition to the affricate /tʃ/, there were three voiceless fricative sibilants, viz., denti-alveolar /s̺/, apical-alveolar /s̺/ and postalveolar /ʃ/. In terms of spelling, the denti-alveolar sound could potentially be represented by any of the letters *z*, *c* (before *e* or *i*) or *ç*; the apical-alveolar sound corresponded to *s* and *ss* and postalveolar /ʃ/ could be written as *x* or as *i/j*. Of these three sounds, the two that are directly relevant to the present discussion are /s̺/ and /s̺/. Examples of words that would, under Alarcos Llorach’s assumptions, have contained them are given in Table 1.

Table 1. The denti-alveolar and apical sibilants of early modern Spanish, according to Alarcos Llorach.

Denti-Alveolar Sibilant (/s̺/)	Apical-Alveolar Sibilant (/s̺/)
<i>caçar</i> ‘to hunt’	<i>esso</i> ‘this’
<i>pozo</i> ‘well’	<i>casa</i> ‘house’

In the modern language, the denti-alveolar member of the above opposition has evolved to /θ/, while postalveolar /ʃ/ has evolved to /x/—processes which [Alarcos Llorach \(1988, p. 53\)](#) attributed to the need ‘to widen the phonetic distance with respect to the apical /s/’ in order to prevent confusion between the relevant sounds.² Less explicitly, but exhibiting a similar concern for structural equilibrium, [Martinet \(1951, p. 151\)](#) linked the fronting of the denti-alveolar sound to the ‘concentration in the sibilant domain’ from which the Spanish phonology suffered, in his view, after the devoicing of the coronal fricatives. In the functionalist model, therefore, the process whereby /θ/ came into existence is seen as being goal-driven, in the sense that it responded to the putative need to maintain a distinction between components of the system. Accordingly, this type of explanation can be described as being broadly teleological.

Modern scholars and researchers have, by and large, adopted the same approach as the one just described, with little or no amendment in terms of the fundamentals. [Penny \(2002, pp. 100–1\)](#), for example, assumes exactly the same post-devoicing scenario as

Alarcos Llorach and he follows the latter author in positing an overtly goal-driven cause for the subsequent evolution of the system. Starting from the premise that there existed a significant potential for confusion between, on the one hand, /s̺/ and /s̠/ (= /s/ in Penny's notation) and, on the other hand, between /s̠/ and /ʃ/, he posits the existence of a need to strengthen the acoustic distinctiveness of each of these contrasts. This was achieved, in his view, 'by exaggerating the contrasts of locus: /s̺/ was moved forwards (away from /s/) and became interdental /θ/, while /ʃ/ was moved backwards (also away from /s/) and became velar /x/' (Penny 2002, p. 101).

The same, centrifugal conceptualization of events can also be found in Ariza 2012, the underlying cause again being the putative need to forestall possible confusion: 'Fricatization and devoicing caused three sibilant phonemes, all fricatives and articulated close to one another, to be at risk of being confused with one another' (Ariza 2012, p. 223).³ Ranson and Quesada (2018, pp. 188–89) posit the same analysis in their introductory textbook on the history of Spanish and it is further assumed by Noll (2021, p. 83) and by Núñez Méndez (2016, p. 75) in their respective discussions of sibilant developments that impacted on Spanish in the Latin American colonies.

In his influential textbook treatment of Spanish phonology, Hualde (2005) largely follows the foregoing account in his overview of sibilant evolution (Hualde 2005, pp. 157–59), positing an early modern sibilant system containing the three voiceless fricatives /s̺/, /s̠/ and /ʃ/, which subsequently 'increased their articulatory distance' (Hualde 2005, p. 158). Interestingly, however, he refrains from explicitly attributing this latter process to a need to prevent confusion. Indeed, he does not appear to touch on the issue of causation. Dworkin (2018, p. 24) similarly refrains from speculating on why /s̺/ evolved to /θ/. However, within his broad conception of multiple changes spreading southwards from the north of the Peninsula, he does appear to assume that the emergence of /θ/ in any given dialect occurred after devoicing had occurred in that dialect. In other words, he follows the standard model in terms of the chronology of the key events.

Setting aside the circumspection, with respect to causation, of Hualde and Dworkin, it seems fair to say that the majority of authors make two key assumptions, viz., (i) that the emergence of Spanish /θ/ postdates the devoicing of the coronal fricatives and (ii) that it responded to an alleged need to increase the articulatory distance between similar-sounding voiceless sibilants. It turns out, however, that these assumptions are based primarily on reconstruction or even conjecture rather than concrete evidence.

As regards the chronology, we can be reasonably certain, on the basis of spelling confusions in the textual record, that devoicing occurred in the sixteenth century (see Menéndez Pidal 1987, p. 115; Ariza 2012, pp. 224–25; Penny 2002, p. 100). On the other hand, the shift from a sibilant articulation of *z*, *ç* and *c^{e,i}* to a non-sibilant one did not in itself trigger any orthographical adjustment (see Section 3, below) and so scholars to date have had no means of establishing when the change took place other than by reference to the ad hoc observations of early modern commentators. However, as Alarcos Llorach (1988, pp. 54–55) forcefully observed, the comments of these *tratadistas* lack both consistency and precision, making them susceptible to multiple interpretations. Above all, what they appear to indicate, in his view, is the co-existence of individual variants—affricates, fricative sibilants and fricative non-sibilants—which, in fact, is exactly what one would expect during a time of linguistic change. Thus, the widely held assumption that the emergence of Spanish /θ/ postdates the devoicing of the coronal fricatives is by no means solidly based on empirical evidence. Indeed, in terms of the usual textual evidence for sound change, specifically changes in the way words are spelled, it is not actually based on any evidence at all.

In this regard, it is worth noting that Menéndez Pidal (1987) implicitly assumed devoicing to have *followed* rather than *preceded* the generalization of the non-sibilant pronunciation of *z*, *ç* and *c^{e,i}*. This can be inferred from his belief that 'already at the beginning of the sixteenth century the interdental, fricative pronunciation θ and *z̧* (= /ð/) was becoming general in many regions of the Peninsula: *plaça, hazer*. The two sounds merged in the XVII

century into a single voiceless one, the voiced sound being lost.’ (Menéndez Pidal 1987, p. 113).⁴ On the basis of new quantitative data, drawn directly from the late medieval and early modern textual record, the present paper will argue that the chronology implicitly assumed by Menéndez Pidal is likely to be the correct one.

Turning to the second of the two above-mentioned core assumptions of the standard model—viz., that the emergence of /θ/ was a reflex of the need to increase the articulatory distance vis-à-vis the apical sibilant /ʃ/—this is discussed in more detail in Section 6.4, below. Suffice it to say, at this stage, that it relies to a significant extent on the conventionally assumed chronology, according to which devoicing preceded the emergence of /θ/. This is because the assumed functionally motivated fronting of /ʃ/ to /θ/ implicitly requires the /ʃ/–/ʃ/ opposition to have had a high functional load (see, e.g., Penny 2002, p. 100), which would not have been the case prior to the loss of the voiced coronal fricatives. For example, prior to devoicing, minimal pairs like *beço* ‘lip’ and *beso* ‘kiss’ would have been distinguished on the basis both of phonation and place of articulation, the *ç* of *beço* being voiceless and the *s* of *beso* being voiced. Thus, if the conventional chronology is poorly motivated empirically, this in turn leaves the associated teleological explanation for the emergence of /θ/ without a solid evidentiary basis.

An additional point to note is that most previous analyses have treated interdentalization as being key to the genesis of Spanish /θ/, articulatory distance concomitantly being conceptualized primarily in terms of place of articulation. However, while the modern Castilian phoneme /θ/ is indeed usually interdental, this is arguably incidental to its auditory properties. English, for example, has a similar sound /θ/, but this appears to alternate between a dental variant in British English and an interdental one in Midwestern and West Coast American English (Ladefoged and Johnson 2014, pp. 12–13), without there being any appreciable difference in its auditory effect. In addition, dental and interdental variants of /θ/ do not appear to contrast phonemically in any language (Ladefoged and Maddieson 1996, p. 143), suggesting that the gestural distinction between them does indeed lack a salient auditory correlate. Therefore, rather than its precise locus of articulation—dental or interdental—the key property of /θ/ in terms of its distinguishability from phonetically adjacent voiceless fricatives such as /s/ is surely the fact that it is a non-sibilant, in the sense that it lacks the latter sound’s ‘high-amplitude, turbulent noise’ (Ladefoged and Johnson 2014, p. 318), itself a by-product of the characteristic groove that is formed at the front of the tongue in sibilant production.

The present paper therefore focuses on *dissibilation*—understood as the converse of assibilation—rather than interdentalization per se. By framing matters in this way, the paper is able to address not just what distinguished /θ/ from its immediate predecessor /ʃ/—the notion which lies at the heart of the standard model—but also in what respects the newer sound was capable of being generated as a variant of the older one. This in turn leads to a revised perspective on the question of causation, in the sense that listener-driven reanalysis turns out to be of potentially far greater significance in the evolution of /θ/ than has previously been suggested.

3. The Linguistic Background

The diachronic pathway that begins with the affricated sibilant /ts/ of early Romance and ends with modern Castilian /θ/ is characterized by a long intermediate stage during which the originally voiceless sound was split into a voiceless and voiced pair, viz., /ts/ and /dz/ (Menéndez Pidal 1987, pp. 112–13). In the Old Spanish orthography, these affricate phonemes were written as *ç* (or *c* before *e* or *i*) and *z*, respectively, except in final and preconsonantal positions, where usually just *z* was used, implying that the /ts/–/dz/ contrast was neutralized in the syllable coda. In the latter case, a following voiced consonant triggered voice assimilation in the affricate and, as regards word-final position, the usual assumption is that the affricate was generally voiceless. Some illustrative examples are given in Table 2, below.

Table 2. Distribution of the letters ζ , $c^{e,i}$ and z in Old Spanish and their phonetic correlates.

Prevocalic (ζ , c or z)	Preconsonantal (z)	Final (z)
<i>cabeça</i> ‘head’ [kaˈbetsa]	<i>diezmo</i> ‘tithe’ [ˈdjedzmo]	<i>foz</i> ‘sickle’ [fots]
<i>tercero</i> ‘third’ [terˈtsero]	<i>lobezno</i> ‘wolf cub’ [loˈbedzno]	<i>assaz</i> ‘enough’ [aˈsats]
<i>pereza</i> ‘sloth’ [peˈredza]	<i>bizconde</i> ‘viscount’ [bitsˈkonde]	<i>ueiez</i> ‘old age’ [βeˈʒets]

Self-evidently, given their modern reflex /θ/, Old Spanish /ts/ and /dz/ must have deaffricated, a process which almost certainly took place in the Late Middle Ages. Crucially, however, the orthography was undisturbed by this change, with all words previously spelled with ζ , $c^{e,i}$ or z continuing to be spelled with these letters, albeit they now stood for the deaffricated reflexes of /ts/ and /dz/.⁵ Moreover, even after the fricative descendants of /ts/ and /dz/ merged in the sixteenth century, although ζ fell into obsolescence, both $c^{e,i}$ and z continued to be used for the surviving, voiceless fricative and they remain in place with that function in modern Spanish. In principle, therefore, the textual record does not offer any direct way of determining when the dissibilated phoneme /θ/ entered the Castilian phonology.

There is, however, an indirect method for ascertaining the relevant chronology, based on the progressive substitution of the letter z for the letter d to denote /d/ in preconsonantal coda position; that is, in clusters consisting of /d/ plus an immediately following non-rhotic consonant. These clusters were secondary or ‘Romance’ consonant groups, in the sense that they had come into existence after the Latin period, specifically as a consequence of the generalized loss of the intertonic vowel (see Penny 2002, pp. 86–87). In some words, the /d/ in such clusters was a direct continuation of the Latin phoneme /d/, while in others the spoken Latin source was /t/, corresponding to orthographic T or, in certain loanwords, to TH.⁶ The first type of derivation is illustrated by *iudgar* ‘to judge’ (< IŪD(I)CĀRE), where the unstressed /e/ (= ĭ) in the antepenult was lost, bringing /d/ into contact with the following velar, which by then had become a voiced sound. It is possible that the /d/ in words of this class was already pronounced as a fricative in late spoken Latin (cf. Menéndez Pidal 1987, pp. 129–30), but even if it was still a plosive at the end of the Latin period, it had almost certainly lenited to [ð] by the time its containing cluster came into existence, as the lenition processes which affected the early Spanish phonology are assumed to have taken place before the generalized loss of the intertonic vowel (Penny 2002, p. 87).

Similar remarks apply to the second of the above-mentioned derivation types, exemplified by the suffix *-adgo* (< -AT(I)CUM) and by the word *bidma* ‘poultice’ (< EPITH(E)MA), although in this type of case two lenition processes must be inferred to have taken place, viz., (i) the voicing of [t] to [d] and (ii) the subsequent fricativization of [d] to [ð]. With respect to the reflexes of the Latin voiceless stops, this is actually quite a general pattern in the history of Spanish, given the language’s well-known phonological rule whereby voiced obstruents must be realized as continuants in intervocalic (and pre-liquid) positions. Penny (2002, p. 76), in fact, specifically alludes to the Latin voiceless stops as instantiating the principle that the output of the voicing process in the lenition chain is capable of becoming the input to the fricativization process.

Until the early modern period, the preconsonantal [ð] just described can be analyzed as being an allophone of the phoneme /d/—a fact reflected in the consistent use of the letter d to represent it. In the modern language, however, it is assigned to the phoneme /θ/, its voiced feature now being analyzed as the product of coda fricative voice assimilation (see, e.g., Hualde 2005, pp. 159–60; Canellada and Kuhlmann Madsen 1987, p. 21). This phonemic reallocation has been accompanied by a change in the spelling, in the sense that preconsonantal coda d has been entirely replaced by z in the inherited lexical stock.⁷ For example, the word now spelled *portazgo* ‘toll’ was invariably spelled as *portadgo* throughout most of the Late Middle Ages, e.g., *non tomen portadgo* ‘they must not levy a toll’ (Fuero

viejo de Alcalá [1223], fol. 54v), *o de tomar los portadgos* ‘or to levy tolls upon them’ (Libro de las leyes [1256–1265], fol. 104r), *que de aqui adelante ninguno non tome portadgo* ‘henceforth no one may levy a toll’ (Ordenamiento de Alcalá [1351–1352], fol. 14v), etc. However, from the fifteenth century onwards, although the spelling with *d* continues for a time—cf. *el portadgo desta çiudad es del enperador* ‘the toll in this city belongs to the emperor’ (Historia del gran Tamorlán [1401–1450], fol. 34r)—its replacement by *z* becomes progressively more frequent: *sin pagar portazgo* ‘without paying a toll’ (Ordenanzas reales [1484], fol. 203v), *obligado a pagar portazgo* ‘obliged to pay tolls’ (Libro primero de las epístolas familiares [1541], fol. 108r), etc. The diachronic trajectory of this change is reconstructed in Section 5, below, on the basis of quantitative data from late medieval and early modern manuscripts and printings.

While the preconsonantal coda allophone of /d/ has undergone phonemic reallocation, as just discussed, the intervocalic and pre-rhotic allophone, as in *boda* ‘wedding’ (< VŌTA) and *padre* (< PATREM), is still assigned to the phoneme /d/ in the modern language. From the phonetic point of view, the intervocalic/pre-rhotic sound is now usually analyzed as being an approximant—in IPA terms [ð̞]—rather than a fricative, due to the apparent absence of audible friction during its articulation (Canellada and Kuhlmann Madsen 1987, p. 37). Moreover, as can be seen from the examples just given, the intervocalic/pre-rhotic spelling has remained unchanged, in the sense that the letter *d* has not been replaced by *z* in this context. The *d* → *z* change has therefore been highly selective, taking place uniformly in preconsonantal coda position but never in intervocalic or pre-rhotic positions. Given this phonologically conditioned selectivity, it seems highly likely that the two allophones of /d/ were already differentiated by the end of the Middle Ages.

It is also important to note that the item which gained a new orthographic representation in this process, viz., preconsonantal coda [ð̞], did not itself undergo any change. Indeed, it survives, unamended, in exactly the same class of words in the modern language: [biðma] *bizma* ‘poultice’, [xuð ʎar] *juzgar* ‘to judge’, [arˈtaðʎo] *hartazgo* ‘surfeit’, etc. Accordingly, what must have changed is the phonological value attaching to the letter *z*, which originally was a sibilant but which, by this time, in order to be a suitable replacement for *d* in the relevant context—preconsonantal coda position—must have become a non-sibilant, specifically a voiced dental one /ð/. Moreover, although the spelling data discussed here refer to preconsonantal coda position, the assumption would have to be that the *z* = /ð/ equivalence applied generally. This follows from the fact that, in order to be spontaneously inclined to substitute *z* for *d* in the relevant context, late medieval and early modern scribes must have perceived a phonetic equivalence between the sound that occurred in that context—viz., [ð̞]—and the sound they associated more generally with *z*. Were that not the case, the replacement of the existing *d* by *z* in the specified context would be entirely mysterious. Accordingly, as was in fact noticed by Martinet (1951), the replacement in question appears to entail that the phonological correlate of *z* was the non-sibilant /ð/, for some speakers at least.⁸

Presumably, the proportion of speakers for whom this was the case was small to begin with but grew over time, because, as is shown in Section 3, below, the *d* → *z* change describes an almost perfect logistic (S-shaped) curve when tracked quantitatively in late medieval manuscripts and early printings. This latter fact also strengthens the case for treating the orthographical *d* → *z* shift as a proxy for an underlying phonological change, given that the parametrized logistic function is widely assumed to be the best model for linguistic change through time (see Altmann et al. 1983; Kroch 1989; Yang 2000; Kauhanen and Walkden 2018). The gradual *d* → *z* shift that can be observed in the textual record thus seems to provide an empirically based means, possibly the only such means, of dating the dissibilation of the reflex of /dz/—a process which, in light of note 5, presumably took the form of a change from denti-alveolar /ʎ/ to /ð/. By the same token, it also provides a means of dating the dissibilation of the reflex of /ts/, given that the reflexes of /dz/ and /ts/ appear to have evolved in parallel. In particular, their merger in the sixteenth century implies that they contrasted minimally, presumably just in terms of phonation and not,

additionally, in terms of sibilance, one item being /ð/ and the other being the voiceless denti-alveolar sibilant /s/.⁹ Indirectly, then, the $d \rightarrow z$ orthographical change gives us an invaluable window onto the chronology of the dissibilation process by which Castilian /θ/ came into existence.

4. Materials and Methods

The present paper tracks the $d \rightarrow z$ change in a corpus of seventy-nine manuscripts and early printings covering the period 1223 to 1542. On the basis of their copy dates (or printing dates, in the case of printed works), these texts have been assigned to the thirteen twenty-year intervals shown in Table 3.¹⁰ The table also shows the unique time value assigned to each period for the purpose of statistical analysis. In each case, this unique value is the mid-point in the date range, rounded to the nearest integer. For example, the time value assigned to the period 1430–1449 is 1440 and the corresponding value assigned to the period 1450–1469 is 1460.

Table 3. Time periods used in the survey.

Time Period	Unique Time Value Assigned
1210–1229	1220
1250–1269	1260
1270–1289	1280
1310–1329	1320
1350–1369	1360
1370–1389	1380
1410–1429	1420
1430–1449	1440
1450–1469	1460
1470–1489	1480
1490–1509	1500
1510–1529	1520
1530–1549	1540

The relatively abundant textual record from the 1410–1429 period onwards enables data points to be distributed at intervals of twenty years, but before that time (with the exception of the Alfonsine period) the scarcity of manuscripts that meet the selection criteria enforces longer gaps between data points. This turns out not to be a problem, however, as the $d \rightarrow z$ change essentially takes place over the course of the fifteenth and early sixteenth centuries, d continuing to be used in over 99% of cases until the end of the fourteenth century.

Full details of the texts in the corpus, including the time periods to which they have been assigned, can be found in Appendix A. With the exception of *La lozana Andaluza* and the three items in the 1530–1549 time period, all of the texts listed were surveyed electronically using the semi-paleographic transcriptions contained in Gago Jover (2011) or in O'Neill (1999). The four items not accessed in this way were surveyed using on-line facsimiles of the relevant manuscripts or early printings indicated in Appendix A.

The basic method involved counting tokens of preconsonantal coda z in which the z replaced an earlier d (e.g., *ye-z-go* 'dwarf elder' instead of *ye-d-go*) and tokens of the more conservative preconsonantal coda d .¹¹ Counts were aggregated by time period, enabling a percentage rate of occurrence, or probability, of z in this context to be calculated for each of the thirteen time periods. The aggregated token counts were then subjected to logistic regression using the `glm` function in R for Mac OS (version 4.4.1).¹² The intercept and slope estimates returned by R were then used to produce a best fit logistic curve, which can be interpreted as modelling the phonological change in question.

5. Results

The numerical results of the survey are given in Table 4, which, for each time period, shows the number of times z occurs in place of preconsonantal coda d , together with the number of times it could have occurred in that role (i.e., $z + d$) and the rate or probability of z occurring in that role (i.e., $z / (z + d)$). Table 5 reports the coefficients from the regression analysis, together with the associated significance estimates (shown as probability values).¹³ Figure 1 provides a visualization of the principal data in Table 4 and also shows the best-fit logistic curve—essentially a logistic trendline—for these data. The logistic curve has been constructed using the intercept and slope coefficients reported in Table 5 as values of the parameters b and a in the parametrized logistic function:

$$f(x) = 1 / (1 + e^{-(ax+b)})$$

Table 4. Replacement of preconsonantal coda d by z (e.g., *juzgo* for *judgo*).

Time Period	z	$z + d$	Probability of z (%)
1210–1229	0	24	0.0
1250–1269	2	1145	0.2
1270–1289	0	236	0.0
1310–1329	0	49	0.0
1350–1369	1	111	0.9
1370–1389	0	2	0.0
1410–1429	8	73	11.0
1430–1449	19	79	24.1
1450–1469	27	105	25.7
1470–1489	258	431	59.9
1490–1509	195	236	82.6
1510–1529	35	36	97.2
1530–1549	105	107	98.1

Table 5. Logistic regression coefficients (returned by the glm function in R).

Coefficient	Estimate	Standard Error	z-Value	Probability
Intercept	−11.431566	0.805484	−14.19	$<2 \times 10^{-16}$ ***
Slope	0.045686	0.003108	14.70	$<2 \times 10^{-16}$ ***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

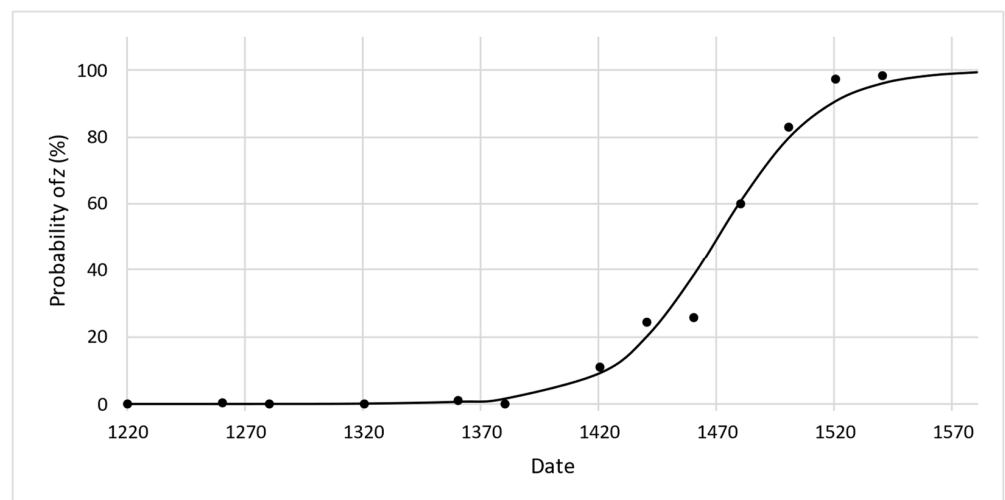


Figure 1. The replacement of preconsonantal coda d by z , together with a best-fit logistic curve. Parameters: −11.431566 (intercept), 0.045686 (slope). p -value: $<2 \times 10^{-16}$ (i.e., almost 0).

The most striking aspect of these results is that the $d \rightarrow z$ shift in the Spanish textual record evolved quantitatively as an almost perfect logistic curve. This can be seen visually in the way that the observed data values (shown as black dots) cluster so closely around the fitted logistic curve in Figure 1. This visual impression is confirmed by the probability value attaching to the regression coefficients, which, at 2×10^{-16} , is extremely close to zero, implying that the odds against getting a match like this by chance would be astronomical. As was noted in Section 3, since the early 1980s, linguistic changes have been widely assumed to evolve as logistic curves and there are now very many empirical studies that confirm that assumption (see [Fruehwald et al. 2009](#) for an example in the domain of phonology). Thus, the finding here that the growth in the use of z (in place of preconsonantal coda d) was logistic is highly significant, suggesting as it does that this orthographical shift was the direct reflection of a linguistic change, specifically a phonological one. If this was not the case—that is, if the relevant diffusion mechanism was non-identical with the diffusion mechanism that characterizes language change—it seems unlikely that the $d \rightarrow z$ shift would have mimicked a genuine linguistic change so perfectly.

6. Discussion

6.1. Dissibilation of Denti-Alveolar /ʒ/ (< /dz/) to /ð/ by the Early 1500s

The quantitative data reported in the previous section thus appear to show a phonological change, enabling the letter z to be used with a value hitherto accorded to d , taking place over the course of the fifteenth century. Given that d did not represent a sibilant, this finding seems likely to be an important one in terms of determining when the sound corresponding to the letter z came to be articulated as a non-sibilant. Indeed, this is arguably the first time that quantitative data have been brought to bear on this question.

Importantly, for the reasons outlined in Section 3, the $d \rightarrow z$ shift does not imply a sound change at the site of the orthographical change. That is to say, the sound originally represented by the letter d in this context, viz., [ð], remained unchanged, even as it came to be denoted by the letter z . What changed was the phonological entity denoted by the letter z , which must have dissibilated to /ð/, thereby becoming indistinguishable from the fricative allophone of /d/ which occurred in preconsonantal coda position. It was this sound change which underlay the replacement of the letter d by z in the relevant context. Thus, the finalization of the $d \rightarrow z$ shift can be interpreted as signaling the completion of a sound change consisting in the dissibilation to /ð/ of the reflex of the medieval affricate /dz/, which is usually assumed to have been a type of /z/, specifically a (laminal) denti-alveolar one /ʒ/ (e.g., see, [Penny 2002](#), p. 99).

According to the results reported in Section 5, the letter d had been almost entirely displaced by z in the relevant context by the 1510–1529 period, at which time the probability of z replacing d was 97.2%. The results thus indicate that denti-alveolar /ʒ/ (< /dz/) had to all intents dissibilated to /ð/ by the early 1500s, even if the theoretical saturation point of 100% is not predicted by the model until about 1570.¹⁴ As was discussed in Section 2, the standard model assumes that devoicing preceded the emergence of /θ/, which implies that Spanish never developed the phoneme /ð/, a voiced counterpart to /θ/. The finding here implies that that assumption may be incorrect. Indeed, the finding is more in line with [Menéndez Pidal's](#) (1987, p. 113) view, according to which two (inter)dental non-sibilant fricatives, /θ/ and /ð/ (the latter denoted by z in his notation), were becoming normal in many regions of the Iberian Peninsula by the beginning of the sixteenth century.¹⁵

6.2. Parallel Emergence of /θ/

While no hard textual evidence is available to confirm the existence in Spanish of the voiceless fricative /θ/ from the early 1500s, it seems unlikely that denti-alveolar /ʒ/ would have evolved to /ð/ by that time without an equivalent dissibilation affecting its voiceless sister phoneme, the denti-alveolar /ʃ/ that is assumed to have evolved from medieval /ts/ (see note 5). Were this not the case, the two sounds would have differed by two auditorily salient features, namely, sibilance and phonation, a circumstance which would

have rendered their merger unlikely. Various data indicate, however, that they did indeed begin to merge from the beginning of the sixteenth century. For example, the well-known remark below, from a work published in 1578 but presumably referring to pre-1540 Spain (1540 being the year its author emigrated to the American colonies), indicates clearly that the merger was already complete in Old Castile by the mid-1500s:

Los de Castilla la Vieja dizen hacer y en Toledo hazer [...]

'The inhabitants of Old Castile say *hacer* but in Toledo they say *hazer*'

(*Del arte en lengua zapoteca*, fol. 68)

This observation is telling because, in the orthography of the time, the letter *c* before *e* or *i* denoted a voiceless dental fricative, as opposed to the voiced one denoted by *z* (i.e., /ð/, according to the findings in Section 5). Thus, the spelling used to depict the pronunciation of the verb *hacer* 'to do' among the Old Castilians is plainly intended to portray a voiceless intervocalic consonant, although previously that same intervocalic consonant was actually voiced and, accordingly, was represented by the letter *z* rather than the letter *c*; that is, *hacer* < *fazer*. The verb *hacer* is evidently being used as an exemplar rather than being referenced as an isolated case, entailing that the voiced dental fricative /ð/ had by the time in question merged in Old Castile with its voiceless counterpart.

Dissibilization of denti-alveolar /ʃ/ to /θ/ by the early 1500s is also suggested by the sharp difference that appears to have existed at that time between French /s/ and the Spanish sound corresponding to the letter *c* as used before *e* or *i*. This difference is revealed rather starkly by Nebrija in his discussion of the pronunciation of the letter *c* in Chapter 9 of *De vi ac potestate litterarum* (1503). Inspired by Quintilian's notion that the letter *c* should be pronounced identically before all vowels, Nebrija laments the fact that, in both Spanish and Italian, *c* was pronounced differently depending on whether it was followed by *a*, *o* or *u* on the one hand, or by *e* or *i* on the other. In the first case, it retained its original velar articulation—cf. Spanish *pecado* 'sin', *flaco* 'thin', Italian *cugino* 'cousin', etc.—whereas in the latter the effects of sound change were apparent: cf. Spanish *tercero* 'third', Italian *città* 'city', in which the *c* no longer represented /k/. While Nebrija does not actually describe the Spanish pronunciation of *c* in this context—no doubt because no other familiar language provided him with a straightforward equivalent—he goes on to state that 'nearly all Frenchmen are no less worthy of ridicule for confusing the sound of this letter [i.e., *c* before *e* or *i*] with the letter *s*'.¹⁶ By implication, therefore, however Spaniards pronounced their *c* before *e* or *i*, what they articulated was not an *s*. Moreover, Nebrija cannot have been referring specifically to a Spanish-style apical or retracted /s/, because the French /s/ corresponding to the letter *c* was the reflex of an earlier denti-alveolar affricate /ts/, which deaffricated in the Middle Ages into a denti-alveolar sibilant. For example, the /s/ = *c* in *forcer* 'to force' evolved from an earlier /ts/—an assibilated reflex of the [tj] that must have arisen in the presumed source, viz., Vulgar Latin *FORTIARE. Thus, the Spanish sound corresponding to *c* before *e* or *i* was not an *s*-like sound, in the broadest possible sense of this concept; that is to say, it was not a sibilant. By a process of elimination, then, it must already have been /θ/.

The analysis advanced here thus assumes a /ts/ > /ʃ/ > /θ/ evolution occurring in parallel to the /dz/ > /ʒ/ > /ð/ change that was indirectly reflected in the orthographic *d* → *z* shift observed in preconsonantal coda position. Moreover, just as the evolution of /ð/ from /ʒ/ was argued to be a general sound change rather than one that was limited to the specific context in which it had an orthographic signature, so the evolution of /θ/ from /ʃ/ should not be assumed to have been conditioned in any way by the preconsonantal coda environment. This point will be particularly relevant in Section 6.4, where reanalysis based on auditory similarity between /ʃ/ and /θ/ is proposed as a causal factor in the genesis of /θ/. Indeed, as an anonymous reviewer observes, the contexts for such reanalysis would not be expected to be limited to the syllable coda.

6.3. Spanish /θ/ Originally Dental Rather than Interdental

The foregoing conclusion does not necessarily mean that the sound in question was already interdental, given that /θ/ can also be dental (see Section 2). Early sixteenth-century Spanish /θ/ was in fact probably apico-dental. Nebrija himself implies this by (i) equating Spanish *ç* with the Hebrew letter samech (*Reglas de ortografía*, fol. 6r) and (ii) attributing a dental articulation to samech: ‘it delivers its sound with the tongue thrust against the base of the upper teeth’ (*De litteris hebraicis*, fol. A. iiii v).¹⁷ In addition, writing at almost the same time, Pedro de Alcalá likened the Arabic letter *ث* (i.e., *thā*³, representing /θ/) to Spanish *ç/c^{e,i}*, but observed that the Arabic sound was articulated with the tongue further forward than was the case in the usual Spanish pronunciation of *ç/c^{e,i}*:

Suena a manera de *c* poniendo el pico de la lengua entre los dientes altos y baxos de manera que suena como pronuncian la *ce* los ceceosos.

‘It sounds like *c* but putting the tip of the tongue between the upper and lower teeth, so that it sounds like the way people with sigmatism pronounce *ce*.’

(*Vocabulista arauigo en letra castellana*, fol. a. iii v)

Guitarte (1992, p. 138) interprets the phrase *los ceceosos* in this context as referring to individuals who exhibited the Andalusian linguistic feature of *ceceo*, in the modern sense, i.e., merger of /s/ and /θ/ in favor of a sound that resembles /θ/ (for a more detailed discussion of this concept, see the second half of Section 6.4, below). While that cannot be ruled out, it seems more likely that Alcalá was referring to an articulation disorder, given that that was how the *cece-* family of terms was generally used at the time, the phonological acceptance evolving later and, in any case, being associated more with the verb *cecear*, together with its deverbal noun *ceceo*, than with the adjectival form *ceceoso*.¹⁸ Indeed, the phenomenon referenced by Alcalá sounds very much like a frontal lisp, the tongue overshooting its target articulation site—the upper teeth, according to Nebrija—and ending up protruding between the front teeth.

Despite Alcalá’s insistence on the foregoing point, the excess fronting of the tongue in the pronunciation of *ç* by the *ceceosos* probably was not auditorily salient. For Nebrija seems to have assumed that *ç* as pronounced by such individuals was identical with the usual pronunciation of *ç* in Spanish. This is apparent in his discussion of Latin *s* in *De litteris hebraicis* (fol. A. iiii v), where he observes that, among the Spanish population, it was actually *los ceceosos* who attained the correct pronunciation, rather than speakers who did not suffer from sigmatism. This rather surprising claim needs to be taken in conjunction with the fact that, owing to some remarks from Martianus Capella, Nebrija had come to believe that Latin *s* had originally been pronounced not like Spanish *s* but like Spanish *ç*. Thus, *los ceceosos*, by not having a sibilant at their disposal, necessarily produced the correct sound when pronouncing Latin *s*, whereas other Spaniards, by deploying a sibilant, were actually committing a mistake. Significantly, given the present discussion, Nebrija does not appear to have regarded the relevant *ceceoso* articulation as being in any way different from the usual pronunciation of Spanish *ç*. Indeed, in the same place, he states that the sole difference between the *ceceosos* and the rest of the population was that ‘we can pronounce both sounds [i.e., *s* and *ç*] while they, owing to an incorrigible defect of the mouth, cannot.’¹⁹

6.4. Causation

As was mentioned in the Introduction, it is commonly assumed that /θ/ came into existence in order to increase the articulatory distance between its immediate ancestor, viz., denti-alveolar /θ̪/, and the retracted or apical /θ̠/ which descended directly from Latin. One obvious objection to this proposal is that if the lack of sufficient articulatory distance between the two sounds was capable of driving them apart, it seems strange that the same factor did not inhibit their convergence in the first place. In other words, the analysis entails that the phonological system moved in two quite contradictory directions, in a relatively short space of time. This cannot be ruled out, but the mere fact that the two sounds did

converge phonetically implies that, linguistically speaking, the situation resulting from this convergence was a viable one. Indeed, a similar contrast, between a laminal /s/ and an apical one, is known to have existed for centuries in Basque, although it has been lost in some varieties (Hualde 1991, p. 10).

More importantly, the analysis in question appears to lack a plausible model of how the alleged need for phonetic divergence would have made itself felt. From the point of view of the late medieval or early modern learner, the articulatory proximity between the two types of /s/ had only two logical outcomes. In the first place, if the sounds could not be properly distinguished, acquisition of the distinction between them would naturally have failed, resulting ultimately in phonemic merger. Alternatively, if the distinction was capable of being successfully acquired, additional phonetic divergence would self-evidently not have been required. Thus, if a need to increase articulatory distance was a factor in the emergence of /θ/, it would necessarily have operated in adult speech. That is to say, assuming that maturation is the primary watershed in language capture, it would have been an L2 effect, where L2 includes adult innovations in the speaker's mother tongue (Postma 2010, p. 270).

Now, unlike in L1 acquisition, in which the grammar is inferred empirically from the primary linguistic data to which the learner is exposed, L2 innovations may be driven by functional pressures, such as a felt need to maximize the phonetic distinguishability of lexical items in order to facilitate communication. A scenario along these lines, implicitly based on the Prague School notion of functional load, seems to be at least tacitly assumed in the teleological model now under consideration. However, if, as is implied by the discussion in Sections 6.1 and 6.2, dissibilation predated devoicing, then at the time the change from /ʃ/ to /θ/ occurred there would have been almost no minimal pairs involving a head-to-head contrast between denti-alveolar /ʃ/ and the apical/retracted sibilant /s/. For example, the commonly cited minimal pair *casa* 'house'–*caça* 'hunt' (see, e.g., Penny 2002, p. 100) would, according to the view developed here, have involved a contrast between some form of /z/ and some form of /s/ rather than between the two specified types of /s/. From this perspective, words like *casa* and *caça* would have been robustly distinguished, on the joint basis of phonation and place of articulation, and there would have been no need for additional reinforcement of the distinction.

Thus, an account of the shift from /ʃ/ to /θ/ that envisages a functional advantage in the increased articulatory distance resulting from the change is not supported by the state of the language at the time. In addition, as was noted above, L1 acquisition is not a plausible locus for a goal-driven process of phonetic divergence. Accordingly, the teleological model of the genesis of /θ/ appears to lack independent motivation; in essence, it seems to confuse outcome with causation.

It is worth noting at this stage that an analogous dissibilation to the one which occurred in Castilian must also have taken place in the other varieties of Romance that exhibit [θ], specifically Franco-Provençal, Venetian and certain dialects of Andalusian (Villena Ponsoda et al. 1994–1995) and Extremaduran (Ariza 1995–1996) Spanish. In those cases, although a denti-alveolar sibilant evolved into a dental non-sibilant, exactly as in Castilian, teleological explanations for this evolution do not appear to have been advanced and the default assumption apparently holds whereby the change in each case reflects spontaneous phonetic evolution. A non-teleological approach, based on listener-driven reanalysis of the phonetic input, may therefore be valid for the emergence of Castilian /θ/, as has in fact been argued for the related phenomenon of sibilant devoicing by Rost Bagudanch (2022), who develops ideas proposed by Ohala (1981, 2012) and Blevins (2004), among others.

In this regard, a key phonetic datum is the one highlighted by Villena Ponsoda et al. (Villena Ponsoda et al. 1994–1995, p. 396, note 12). They observe that the auditory distinction between /θ/ and denti-alveolar /ʃ/ (in their terms 's/ dorsodental') is extremely precarious, given Quilis's (1981, p. 236) experimental finding that 'as the place of articulation advances towards the dental area, stridency gives way to dullness, which becomes apparent in the spectrogram of the laminal denti-alveolar [s] . . . This dullness correlates

with a more regular distribution of the frequency bands, which results in spectrograms similar to those of [θ].²⁰ The sound [θ] and the denti-alveolar type of [s] should thus be seen as two idealized extremes on a single continuum, with marginal adjustments in articulation delivering sounds that a listener may not be able to categorize with certainty.

In practical terms, this indeterminacy has long been known to commentators on Andalusian *ceceo*, the phenomenon whereby the merger of the medieval Spanish phonemes /ts/, /dz/, /s/ and /z/ has resulted in a variable sound that oscillates between a sibilant articulated at the outer boundary of sibilance and the genuine non-sibilant [θ]. For example, while investigating *ceceo* as part of their famous survey of Andalusian Spanish, Navarro Tomás et al. (1933, p. 270) reported finding it almost impossible in some cases to determine whether a speaker was articulating [θ] or some form of the ‘s predorsal’, i.e., denti-alveolar /ʃ/. In the same vein, Dalbor (1980, p. 9), alludes to a ‘fuzzy’ or ‘imprecise’ Andalusian /s/, which he assesses as being auditorily very similar to the dental variety of /θ/. Similarly, in their survey of *ceceo* in Malaga city, Villena Ponsoda et al. (Villena Ponsoda et al. 1994–1995, p. 393, note 2) report finding ‘abundant indeterminate cases’ (‘abundantes casos dudosos’) in which the researcher could not decide between /θ/ and /s/. In fact, it is precisely this potential for convergence between the denti-alveolar type of [s] and the non-sibilant [θ] which must have enabled the Andalusian merger of the medieval sibilants /ts/, /dz/, /s/ and /z/ to take the two apparently contrasting forms that are observable in the modern era, viz., the *ceceo* phenomenon, based in principle on a non-sibilant, as described above, and *seseo*, in which the unique reflex is a voiceless denti-alveolar sibilant. The lesson that can be drawn from this dual historical outcome, together with the tendency reported by fieldworkers for the two sounds to become indistinguishable, must surely be that the auditory threshold between denti-alveolar [ʃ] and [θ] is in practice highly permeable.

This finding has a direct bearing on the situation in late medieval Castilian, the phonology of which is assumed to have contained a denti-alveolar type of /s/ which would in due course evolve into /θ/. Given the foregoing discussion, it can be assumed that slight advances in the place of articulation of this denti-alveolar /ʃ/, randomly produced by speakers as a by-product of the natural variability of human speech production, would have sufficed to make tokens of this phoneme sound auditorily similar or identical to the non-sibilant [θ]. Adult speakers who had already acquired /ʃ/ as part of their grammar would no doubt have been unaffected by this convergence, given the array of lexical and syntactic cues that can compensate for imperfect phonetic realization. L1 learners, in contrast, still engaged in the process of inferring their phonology from the primary linguistic data, would have been in a quite different situation. They, of necessity, rely entirely on their auditory perception in order to build hypotheses about their phonology; consequently, nothing would have prevented them from analyzing any slightly fronted tokens of /ʃ/ as [θ]. It therefore seems entirely plausible to suppose that some late medieval learners did execute such an analysis, subsequently incorporating a reanalyzed phoneme /θ/ into their adult grammars, from where it became available to diffuse across the speech community. As was discussed in Section 6.3, this /θ/ was probably dental to begin with, advancing only later to the interdental locus it exhibits in the modern language.

7. Conclusions

This article has proposed a revised model for the genesis of /θ/ in Castilian Spanish. According to the view advanced, the deaffricated reflexes of /dz/ and /ts/ dissibilated, to /ð/ and /θ/, respectively, by the early 1500s and subsequently merged, the latter process reaching completion in Old Castile by about 1540. The major evidence for this chronology comes from the orthographical *d* → *z* change in preconsonantal coda position, which took place over the course of the 1400s and reached effective completion by the period 1510–1529. At the phonological level, this change diagnoses an underlying process of dissibilation, whereby denti-alveolar /ʃ/ (< /dz/) evolved to /ð/. The fact that /ð/ merged with a voiceless fricative shortly afterwards implies that this latter, voiceless item must have been the non-sibilant /θ/. Had it still been a sibilant, specifically denti-alveolar /ʃ/, it seems

highly unlikely that merger would have occurred, as the two sounds would have differed in terms of two salient features. Nebrija's implicit assumption that the sound corresponding to the Spanish letter *c* when it occurred before *e* or *i* was not a sibilant confirms the early existence of /θ/.

The proposed chronology, whereby dissibilation preceded devoicing, implies that vanishingly few minimal pairs would ever have been distinguished on the basis of a direct contrast between the denti-alveolar /ʃ/ (< /ts/) and the apical /s/ that descended from Latin /s/. Accordingly, the functional argument for a goal-driven shift from /ʃ/ to /θ/ is left with little in the way of solid empirical motivation. Conversely, the auditory threshold between these two sounds is known to be highly permeable, a circumstance which is strikingly manifested by the Andalusian *seseo-ceceo* dichotomy itself and by the indeterminate nature of the phonetic entity that characterizes the *ceceo* dialects. Given this permeability, it was argued that the dissibilation of /ʃ/ was a reflex of its auditory proximity to /θ/, with late medieval learners reanalyzing marginally fronted instances of /ʃ/ as /θ/ and ultimately acquiring the latter sound in place of the former. The revised model thus eschews the traditional teleological approach to the genesis of /θ/ and in that sense exhibits a certain similarity to the stance adopted by [Rost Bagudanch \(2022\)](#) in her treatment of the historically related phenomenon of sibilant devoicing.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data reported in this study are openly available within the article itself.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Texts used in the corpus, grouped by time period

1210–1229

Fuero viejo de Alcalá (Alcalá de Henares, AMA (H) F.V.A)

1250–1269

Cánones de Albateni (Paris: Arsenal 8322)

Judizios de las estrellas (BNE, MSS/3065)

Lapidario (Escorial, h-I-15)

Libro de las animalias que cazan (BNE, RES/270)

Libro de las cruces (BNE, MSS/9294)

Libro de las leyes (British Library, Add. 20787)

Picatrix (Vatican Reg. Lat. 1283)

Tablas de Zarquiel (Paris: Arsenal 8322)

1270–1289

Estoria de España I (Escorial, Y-I-2)

Formas e imágenes (Escorial, h-I-16)

Fuero Juzgo (HSA, B2567)

General estoria I (BNE, MSS/816)

General estoria IV (Vatican Urb lat 539)

Libro de ajedrez, dados y tablas (Escorial, T-I-6)

Libro del cuadrante señoero (Paris: Arsenal 8322)

Libros del saber de astronomía (Madrid: Biblioteca Universitaria Complutense 156)

Poridat de las poridades (Escorial, L-III-2)

1310–1329

Estoria de España II (Escorial, X-I-4)

- General estoria VI (Toledo: Catedral 43-20)
- 1350–1369
- Libro de la montería (Escorial, Y.II.19)
- Ordenamiento de Alcalá (BNE, vitrina 15-7)
- 1370–1389
- Visita y consejo de médicos (BNE, MSS/18052)
- 1410–1429
- Crónica de 1344 I (Biblioteca Francisco de Zabáburu y Basabe, 11-109)
- Historia del gran Tamorlán (BNE, MSS/9218)
- Moreh Nevukim; Mostrador y enseñador de los turbados (BNE, MSS/10289)
- Sermones contra los judíos y moros (Soria: Biblioteca pública del estado 25-H)
- 1430–1449
- Axioco (BNF: Espagnol 458)
- Doce trabajos de Hércules (BNE, MSS/27)
- Libro de las donas (Escorial, h-III-20)
- Libro de los ejemplos por A.B.C. (BNE, MSS/1182)
- Tratado de la reformation de la ánima (BNF: Espagnol 458)
- 1450–1469
- Cancionero de Salvá (BNF, Espagnol 510)
- Conjuración de Catilina (Escorial, g.III.11)
- Corbacho (Escorial, h-III-10)
- Espejo de medicina (BNE, MSS/3384)
- Invencionario (BNE, MSS/9219)
- Morales de Ovidio (BNE, MSS/10144)
- 1470–1489
- Arte de bien morir-Breve confesionario (Escorial, 32-V-19⁴)
- Atalaya de las Crónicas (British Library, Egerton: MS 287)
- Caída de príncipes (HSA B1196)
- Compilación de las batallas campales (Murcia: Lope de la Roca, 1487)
- Crónica de España (Seville: Alfonso del Puerto, 1482)
- Escritura de cómo y por qué razón no se debe dividir, partir ni enajenar (Murcia: Lope de la Roca, 1487)
- Esopete ystoriado (John Rylands Library 19562)
- Historia de la linda Melosina (Bibliothèque royale de Belgique: Inc. B 840)
- Letra sobre los matrimonios y casamientos entre los reyes de Castilla (Murcia: Lope de la Roca, 1487)
- Ordenanzas reales (Huete: Álvaro de Castro, 1484)
- Secretos de la medicina (Madrid: Real Biblioteca II/3063)
- Tratado de la adivinanza (BNE, MSS/6401)
- Tratado de las fiebres (Escorial, M-I-28)
- Valerio de las historias escolásticas y de España (Murcia: Lope de la Roca, 1487)
- 1490–1509
- Cárcel de amor (Seville: Paulus von Köln et al., 1492)
- Claros varones de Castilla (Seville: Stanislaw Polak, 1500-04-24)
- Compendio de medicina (Biblioteca Universitaria de Salamanca, 2262)
- Cura de la piedra (Toledo: Pedro Hagenbach, 1498)
- Ejemplario contra los engaños y peligros del mundo (Zaragoza: Pablo Hurus, 1493)
- Enrique fi de Oliva (Seville: Johann Pegnitzer von Nürnberg et al., 1498)
- Generaciones y semblanzas (Fundación Lázaro Galdiano, IB 14498)
- Gramática castellana (Salamanca: Juan de Porras, 1492)
- Las pronósticas (Seville: Meinhard Ungut and Stanislaw Polak, 1495)

Libro de albeitería (Zaragoza: Pablo Hurus, 1499)
 Libro Llamado Infancia Salvatoris (Burgos: Juan de Burgos, ca. 1493)
 Lilio de medicina (Seville: Meinhard Ungut and Stanislaw Polak, 1495)
 Menor daño de medicina (Escorial, b-IV-34)
 Oliveros de Castilla (New York, HSA)
 Propiedades de las cosas (Toulouse: Heinrich Meyer, 1494)
 Sumario de la medicina (Salamanca: Juan de Porras, 1498)
 Tratado de la fisionomía en breve suma contenida (Zaragoza: Pablo Hurus, 1494)
 Tratado de la peste (Zaragoza: Pablo Hurus, 1494)
 Tratado en defensa de virtuosas mujeres (BNE, MSS/1341)
 Triunfo de amor (BNE, MSS/22019)

1510–1529

Arnalte y Lucenda (Biblioteca Trivulziana Cod. Triv. 940)
 Cuaderno de las leyes nuevas de la hermandad (Seville: Jacobo Cromberger, ca. 1511)
 La lozana andaluza (Venice, 1528)
 Obra de agricultura de Gabriel Alonso de Herrera (Alcalá de Henares: Brocar, 1513)

1530–1549

Diálogo de la lengua (BNE, MSS/8629)
 Libro primero de las epístolas familiares (Valladolid: Juan de Villaquirán, 1541)
 Segunda Celestina (Venice: Sabio, 1536)

Appendix B

Primary sources cited but not used in the corpus

Alcalá, Pedro de. 1505. *Vocabulista arauigo en letra castellana*. Granada: Juan Varela.
 Córdoba, Fray Juan de. 1578. *Del arte en lengua zapoteca*. México: Pedro Balli.
 Nebrija, Antonio de. 1503. *De vi ac potestate litterarum*. Salamanca: Juan Gysser.
 Nebrija, Antonio de. 1515. *De litteris hebraicis cum quibusdam annotationibus in scripturam sacram*. Alcalá de Henares: Brocar.
 Nebrija, Antonio de. 1517. *Reglas de ortografía en la lengua castellana*. Alcalá de Henares: Brocar.

Notes

- ¹ In the general case, the [tj] sequence arose in words like PUTEUM ‘well’ and VITIUM ‘vice’ once the unstressed front vowel had ceased to be syllabic. It also evolved, in many dialects of spoken Latin, from the palatals [cj] (BRACHIUM ‘arm’) and [cʲ] (VICINUM ‘neighbor’). Where the latter development occurred, the relevant modern language usually has the same sound (or type of sound) in the reflexes of words like BRACHIUM and VICINUM as in the reflexes of words like PUTEUM and VITIUM (cf. Castilian Spanish *brazo*, *vecino*, *pozo* and *vicio/vezo*, all with /θ/).
- ² ‘El desplazamiento de ambos fonemas reside en el intento de ampliar al margen de seguridad respecto a la ápico-alveolar /s/’.
- ³ ‘La fricativización y el ensordecimiento produjeron que tres fonemas sibilantes, los tres fricativos y de articulación muy próxima, corriesen el peligro de confundirse’.
- ⁴ ‘A comienzos del siglo XVI ya se generalizaba en muchas regiones de la Península la pronunciación interdental, simplemente fricativa θ y z: *plaçā, hazer*. Ambos sonidos se confundieron a partir del siglo XVII en un solo sordo, perdiéndose el sonoro.’
- ⁵ Little can be known with certainty about these reflexes, but the prevailing assumption in the literature (Lapesa 1981; Penny 2002; Dworkin 2018, etc.) is that they were laminal denti-alveolar sibilants, transcribable as /ʃ/ and /ʒ/, respectively.
- ⁶ The Latin digraph TH was presumably articulated as [tʰ] by at least some, no doubt highly educated, speakers. However, the aspirated articulation does not seem to have been normal among the general populace, at least not in northern central Iberia, where Spanish emerged. This can be inferred from the fact that, in the history of Spanish, the phonological correlate of Latin TH evolves in exactly the same way as does the phonological correlate of Latin T. Compare, for example, CATHEDRAM > *cadera* ‘hip’ with CATENAM > *cadena* ‘chain’.
- ⁷ This process of orthographical substitution did not apply to cultismos like *admirar* and *advertir*, where the preconsonantal *d* was directly modelled on the Latin spelling.
- ⁸ The phonemic status of this /ð/ stems from the fact that the letter *z* represented a discrete phoneme. Moreover, despite the encroachment of *z* into an orthographic space previously occupied by *d*, the sounds represented by these letters must have continued to contrast phonemically. For example, under the assumptions advanced here, minimal pairs such as *lazo*

'loop/noose/trap' and *lado* 'side' would have been distinguished on the basis of a /ð/–/d/ contrast, which, following the devoicing of the coronal fricatives, is now a /θ/–/d/ contrast.

9 That the two sounds differed minimally at the time is independently confirmed in Juan de Valdés's *Diálogo de la lengua* (c. 1535), where it is noted that applying the cedilla to the letter *c* 'la haze sonar cassy como .z.' ('makes it sound almost like z') (MSS/8629, fol. 60r).

10 Where a manuscript's known or estimated copy date is a range rather than a specific date, it has been assigned on the basis of the mid-point in the range. For example, the PhiloBiblon database (Faulhaber 1997–) gives the range 1436–1450 as the copy date for the manuscript of the *Libro de los ejemplos por A.B.C.* used in this study. The mid-point in this range is 1443, so the text is assigned to the 1430–1449 period.

11 This excludes, therefore, the *z* in words like *diezmo* 'tithe', where it corresponds to a (palatalized) Latin velar, viz., the /k/ (= [cʰ]) in DECIMUM. Conversely, the *d* of *cultismos* like *admirar* and *advertir* is also excluded, as this *d* was never replaced by *z*.

12 If the data are arranged as a three-column matrix in a text file named 'dataset.txt', with headers 'date', 'yes' (= count of *z*) and 'no' (= count of *d*), the relevant command lines are the ones shown below:

```
> mydata = read.table("dataset.txt", header=TRUE, sep="\t",)
> mylogit <- glm(cbind(yes, no) ~ date, data = mydata, family = "binomial")
> summary(mylogit)
```

13 The coefficients in Table 5 are based on 1220 (the mid-point in the first twenty-year period) being treated as year zero, with all subsequent time values recalibrated accordingly (i.e., 1260 becomes year 40, 1280 become year 60, etc.). The choice of year zero is mathematically arbitrary and does not affect the regression analysis per se (Kroch 1989, p. 225), because the values of the logistic function go from minus infinity to plus infinity. However, for the purpose of plotting the fitted logistic curve on a line chart, as in Figure 1, the values assigned to *x* in the logistic equation must match the time value recalibration that stems from the particular choice of year zero.

14 When represented as a logistic curve, the initial and final years of a linguistic change are, by definition, statistical outliers and hence should not be thought of as being integral components of the change event. The real-world analogue of this is that speakers experiencing the change would be unlikely to perceive a highly infrequent variant, be it an innovative one at the start of the change or a conservative one at the end, as evidence of genuine variation.

15 Conceivably, this early modern /ð/ is also the source of the /ð/, notated as *d*, which occurs in Chinato, the nearly extinct dialect of Extremaduran spoken in Malpartida de Plasencia (see Ariza 1995–1996). It should be noted, however, that in that dialect, /ð/ occurs not just in words which in Old Spanish had /dz/ but also in those which had /z/ (< Latin /s/); for example, *didil* 'to say' (Old Spanish: *dezir* [de dzir]) and *cada* 'house' (Old Spanish: *casa* [kaza]).

16 'Neque sunt ridendi minus fere omnes galli qui huius litterae sonum cum s. littera confundunt.'

17 '[...] ad supernorum dentium radices lingua illisa sonum reddit.'

18 This can be seen, for example, in the differential manner in which *ceceo* and *ceceoso* are defined in the Real Academia Española's *Diccionario de Autoridades* (1726). The first term is defined there purely in linguistic terms, viz., as 'la pronunciacion de la persona que trueca la S en C' ('the pronunciation of someone who merges S with C') and the entry includes a reference to a verse from Quevedo that mocks the 'cecéos' of Andalusians. In contrast, a *ceceoso* is defined as someone who suffers from a natural disorder, albeit one which has a linguistic consequence: 'el que naturalmente y sin poderlo remediar muda en las palabras la pronunciacion de la S en C' ('that person who naturally and without being able to remedy it changes, within words, the pronunciation of S into C').

19 Sed nos illos hac una in re superamus: quod utramque vocem possumus efferre: illi vero inemendabili oris pravitate non possunt.'

20 'A medida que el lugar de articulación va avanzando y se sitúa en la proximidad dental, la estridencia va disminuyendo, dejando paso a la cualidad de mate, que se hace patente en el espectro de la [s] predorsodentoalveolar [...] La característica mate lleva consigo una distribución más regular de las regiones de frecuencias, distribución que origina unos espectros semejantes a los de [θ].'

References

- Alarco Llorach, Emilio. 1988. De nuevo sobre los cambios fonéticos del siglo XVI. In *Actas del I Congreso Internacional de Historia de la Lengua Española*. Edited by Manuel Ariza, Antonio Salvador and Antonio Vindas. Madrid: Arco Libros, pp. 47–59.
- Alarco Llorach, Emilio. 1991. *Fonología española*, 4th ed. Madrid: Gredos.
- Altmann, Gabriel, Haro Buttlar, Walter Rott, and Udo Strauss. 1983. A law of change in language. In *Historical Linguistics*. Edited by Barron Brainerd. Bochum: Studienverlag Dr. N. Brockmeyer, pp. 104–15.
- Ariza, Manuel. 1995–1996. De nuevo sobre Serradilla y el chinato. *Cauce. Revista de Filología y su Didáctica* 18–19: 689–701.
- Ariza, Manuel. 2012. *Fonología y fonética históricas del español*. Madrid: Arco Libros.
- Blevins, Juliette. 2004. *Evolutionary Phonology. The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Canellada, María Josefa, and John Kuhlmann Madsen. 1987. *Pronunciación del Español. Lengua Hablada y Literaria*. Madrid: Castalia.
- Dalbor, John B. 1980. Observations on present-day *seseo* and *ceceo* in Southern Spain. *Hispania* 63: 5–19. [CrossRef]

- Dworkin, Steven N. 2018. *A Guide to Old Spanish*. Oxford: Oxford University Press.
- Faulhaber, Charles B. 1997–. PhiloBiblon. Available online: <https://bancroft.berkeley.edu/philobiblon/index.html> (accessed on 11 November 2021).
- Fruehwald, Josef, Jonathan Gress-Wright, and Joel C. Wallenberg. 2009. Phonological change: The constant rate effect. In *North East Linguistic Society (NELS) 40*. Amherst: GLSA.
- Gago Jover, Francisco. 2011. Digital Library of Old Spanish Texts. Hispanic Seminary of Medieval Studies. Available online: <http://www.hispanicseminary.org/textconc-en.htm> (accessed on 11 November 2021).
- Hinzelin, Marc-Olivier. 2018. Contact-induced change in Franco-provençal phonological systems caused by standard French. *International Journal of the Sociology of Language* 249: 49–70.
- Guitarte, Guillermo. 1992. Cecear y palabras afines. In *Actas del II Congreso Internacional de Historia de la Lengua Española Tomo I*. Edited by Manuel Ariza, Rafael Cano, Josefa Mendoza and Antonio Narbona. Madrid: Pabellón de España, pp. 127–64.
- Hualde, José Ignacio. 1991. *Basque Phonology*. London: Routledge.
- Hualde, José Ignacio. 2005. *The Sounds of Spanish*. Cambridge: Cambridge University Press.
- Kauhanen, Henri, and George Walkden. 2018. Deriving the constant rate effect. *Natural Language and Linguistic Theory* 36: 483–521. [CrossRef]
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199–244. [CrossRef]
- Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Ladefoged, Peter, and Keith Johnson. 2014. *A Course in Phonetics*, 7th ed. Boston: Cengage.
- Lapesa, Rafael. 1981. *Historia de la lengua española*, 9th ed. Madrid: Gredos.
- Martinet, André. 1951. The Unvoicing of Old Spanish Sibilants. *Philology* 5: 133–56.
- Menéndez Pidal, Ramón. 1987. *Manual de gramática histórica española*, 19th ed. Madrid: Espasa Calpe.
- Navarro Tomás, Tomás, Aurelio M Espinosa, and L. Rodríguez Castellano. 1933. La frontera del andaluz. *Revista de Filología Española* 20: 225–77.
- Noll, Volker. 2021. The Emergence of Latin American Spanish. In *English and Spanish: World Languages in Interaction*. Edited by Danae Perez, Marianne Hundt, Johannes Kabatek and Daniel Schreier. Cambridge: Cambridge University Press, pp. 76–91. [CrossRef]
- Núñez Méndez, Eva. 2016. A diachronic approach to the Old Spanish sibilant merger and its impact on trans-Atlantic Spanish (part I). *Revista de Estudios Hispánicos* 3: 59–98.
- Ohala, John. 1981. The listener as a source of sound change. In *Papers from the Parasession on Language and Behavior*. Edited by Carrie S. Masek, Roberta A. Hendrick and Mary Frances Miller. Chicago: Chicago Linguistic Society, pp. 178–203.
- Ohala, John. 2012. The listener as a source of sound change. An update. In *The Initiation of Sound Change: Perception, Production, and Social Factors*. Edited by Maria-Josep Solé and Daniel Recasens. Amsterdam: John Benjamins, pp. 21–36.
- O'Neill, John. 1999. *Electronic Texts and Concordances of the Madison Corpus of Early Spanish Manuscripts and Printings*. CD-ROM. Madison: Hispanic Seminary of Medieval Studies.
- Penny, Ralph. 2002. *A History of the Spanish Language*, 2nd ed. Cambridge: Cambridge University Press.
- Postma, Gertjan. 2010. The impact of failed changes. In *Continuity and Change in Grammar*. Anne Breitbarth, Christopher Lucas, Sheila Watts and David Willis. Amsterdam: John Benjamins, pp. 269–302.
- Quilis, Antonio. 1981. *Fonética Acústica de la Lengua Española*. Madrid: Gredos.
- Ranson, Diana L., and Margaret Lubbers Quesada. 2018. *The History of Spanish: A Student's Introduction*. Cambridge: Cambridge University Press.
- Rost Bagudanch, Assumpció. 2022. More on Sibilant Devoicing in Spanish Diachrony: An Initial Phonetic Approach. *Languages* 7: 27. [CrossRef]
- Villena Ponsoda, Juan Andrés, José María Sánchez Sáez, and Antonio Manuel Ávila Muñoz. 1994–1995. Modelos probabilísticos multinomiales para el estudio del ceceo, seseo y distinción de /s/ y /θ/. Datos de la ciudad de Málaga. *Estudios de Lingüística Universidad de Alicante* 10: 391–436. [CrossRef]
- Widdison, Kirk A. 1987. 16th Century Spanish sibilant reordering: Reasons for divergence. *Deseret Language and Linguistic Society Symposium* 13: 63–70.
- Yang, Charles D. 2000. Internal and external forces in language change. *Language Variation and Change* 12: 231–50. [CrossRef]
- Zamboni, Alberto. 1974. *Veneto (Profilo dei Dialetti Italiani, 5)*. Pisa: Pacini.