UNIVERSITAT DE BARCELONA

Estudios de Fonética Experimental
Journal of Experimental Phonetics
ISSN: 1575-5533 · ISSN-e: 2385-3573

EFE
Estudios de Fonética Experimental
Journal of Experimental Phonetics

Acoustic phonetics: Notes on the evolution of /ti̯, ki̯/ yod contexts in Spanish

Assumpció Rost Bagudanch [1] 0000-0001-8133-881X

[1] Universitat de les Illes Balears (Spain)

DOI: 10.1344/efe-2023-32-169-194

Corresponding address: assumpcio.rost@uib.es

Received: 27/03/2023 Accepted: 12/07/2023 Published: 28/11/2023

Rost Bagudanch, A. (2023). Acoustic phonetics: notes on the evolution of /ti̯, ki̯/ yod contexts in Spanish. *Estudios de Fonética Experimental*, *32*, 169–194. https://doi.org/10.1344/efe-2023-32-163-194

Abstract

The explanation for the evolution of /ti̯/ and /ki̯/ yod contexts has been based on articulatory reasons. However, sound change has an important auditory and acoustic basis. This paper deals with two main topics: (a) it examines the confluence of the phonetic realisations of dental and velar stops in /ti̯/ and /ki̯/, and (b) it seeks to determine whether voicing occurs after the assibilation process. The results show that the allophones for /t/ and /k/ display acoustic similarities regarding point of articulation cues, which could explain misperceptions. Voicing also seems to be very frequent in these phonetic realisations, which points to a process simultaneous with assibilation.

Keywords

historical phonetics; /ti̯, ki̯/ yod contexts; acoustic phonetics; assibilation; sibilant voicing

Fonètica acústica: Apunts sobre l’evolució de iod 1a en espanyol

RESUM

L’evolució dels contextos de iod 1a s’ha explicat tradicionalment des de la fonètica articulatòria. No obstant això, els processos de canvi fònic tenen un component acústic i perceptiu important. Aquest treball vol ser una primera aproximació a aquest fenomen des de la fonètica acústica. Aquest treball aborda dos temes principals: (a) examina la confluència de les realitzacions fonètiques de les oclusives dentals i velars en /ti̯/ i /ki̯/, i (b) busca determinar si la seva sonorització és posterior a l’assibilació. Els resultats mostren que els al·lòfons de /t/ i /k/ mostren similituds acústiques relacionades amb el punt d’articulació, fet que podria explicar la confusió. A més, les dades assenyalen que la sonorització és molt freqüent, fet que apunta a un procés simultani amb l’assibilació.

MOTS CLAU

fonètica històrica; iod 1a; fonètica acústica; assibilació; sonorització de sibilants

1. Introduction

The explanation of the evolution of /ti̯/ and /ki̯/ yod contexts has traditionally been based on articulatory studies, some more impressionistic, others with a solid empirical basis, especially regarding variation and diachrony in other languages, such as Catalan (Recasens, 2011) or German (Hall et al., 2006), to cite two examples. However, there are relatively few studies dealing with the acoustic aspect of change and there is a lack of work on the perceptual part of the process (Chang et al., 2001, and Guion, 1998, are two exceptions). This paper proposes a first experimental approach to the study of this yod process in Spanish from the perspective of acoustic phonetics.

On this occasion, we are interested in explaining how the assibilation of /t/ and /k/ occurs and the convergence of solutions between two sequences that are different in origin. We will also deal with their voicing, although more briefly. The simplification of the diphthong and the final result as an interdental fricative will be left for later studies.[1]

As indicated, the purpose of this paper is to examine the phonetic reasons underlying these changes by means of an experimental approach.

Logically, in this case, it is not possible to begin with original sources, since ancient documentation is written, not oral, and, although writing provides very important clues about oral forms, it does not allow for a direct analysis of the phonetic phenomena that take place in the contexts under study. The solution, when it comes to sound change, lies in using current speech samples in order to examine diachrony through synchrony. Indeed, the literature has already reported on the close relationship between synchronic variation phenomena and the description of historical changes, and how the analysis of the former can shed light on the latter (De Vogelaer & Seiler, 2012, p. 3; Harrington, 2012; Hinskens, 2021, p. 36). This approach is being adopted here, since the aim is to start from present-day spontaneous speech in order to examine contexts analogous to those that have triggered changes in present-day Spanish.

The objectives to be achieved are multiple. Firstly, we want to acoustically analyse sequences of [t̪i̯] and [ki̯] in spontaneous speech to establish whether dental and velar consonants show allophonic variation or whether they are realised homogeneously. If variation is detected, the phonetic nature of the allophones documented needs to be determined. This will require us to (a) describe the acoustic features of the different solutions and (b) analyse whether the acoustic parameters usually associated with the point of articulation in stops point to a possible convergence between these realisations. Furthermore, we aim to determine the degree of voicing of these acoustic manifestations and whether, if this process occurs, it is related to the preceding context, as stated in the bibliography.

Thus, two main hypotheses are assumed. Firstly, it is postulated that /t/ and /k/ followed by a diphthong will show phonetic variation, which can be partly explained by a coarticulatory process triggered by the palatal semivowel. Among the possible allophones, there should be realisations with a certain level of assibilation that could coincide in both types of sequence, as well as voiced solutions, especially in intervocalic contexts. Secondly, we argue for the existence of coincidental allophones in both sequences, especially with aspiration or friction, which will make it possible to show that the confluence between the two types of sequence is due to common phonetic tendencies related to the process of palatalisation and subsequent assibilation. Likewise, it is considered that the existence of acoustic manifestations with different degrees of voicing will make it possible to show that this process must be simultaneous, not a later phase.

The results should illustrate diachronic evolution, which can be explained as a process of progressive

[1] Cf. García Santos (2001), for an experimental acoustic approach to the emergence of /θ/.

assibilation under the influence of the semivowel.

The paper is organised as follows: Section 2 briefly summarises the descriptions of the historical process concerned (§2.1), reviews the concepts of palatalisation and assibilation as applied to the contexts /ti̯/ and /ki̯/ (§2.2), details the phonetic characteristics of the sounds involved in the change (§2.3) and gives an account of the mechanisms of voicing (§2.4). Section 3 describes the experimental design. The results are reported in Section 4: §4.1 presents the results concerning allophonic variation from the perspective of assibilation and point of articulation in the sequences /ti̯/ and /ki̯/, while §4.2 gives the results on consonant voicing. Section 5 is devoted to the discussion of these results, and brief conclusions are given in Section 6.

2. Some fundamental issues

2.1. The evolution of /ti̯, ki̯/ yod contexts and the explanation of sound change

Menéndez Pidal coined the term *yod* to refer to a whole series of diachronic processes of palatalisation which usually involved the appearance of a palatal semivowel in the evolution of that sequence. This semivowel could affect not only the consonant preceding it, but also the vowels (Menéndez Pidal, 1904/1985, pp. 45–50). Among these occurrences, /t/ and /k/ followed by a palatal semivowel are the earliest: already in Vulgar Latin (2nd century AD), they would have been produced as a dentoalveolar affricate, perfectly established in medieval times (Ariza, 2012, pp. 145–146, 155; Lapesa, 1981, pp. 79–80; Lloyd, 1993, pp. 220–221; Quilis, 2005, pp. 114–119). In general terms, it is assumed that, in order to reach the affricate solution, the stops are first palatalised, with different solutions for each of them. This part of the process is common to the Romance languages (Loporcaro, 2011, pp. 143–144; Repetti, 2016, p. 659; Zampaulo, 2019, pp. 94–95).[2]

These distinct evolutions for the two contexts would eventually converge: they were sounds that could be confused relatively easily (Lloyd, 1993, p. 221; Ariza, 2012, p. 28). In fact, Quilis (2005, pp. 115–116) explicitly refers to the existence of variation in the evolution of the original sequences, producing sibilants in both cases. These results would have been similar enough to give rise to frequent confusion. This author links their final convergence to the evolution of contexts of /k, g/ + palatal vowel, which appear around the 5th century, also palatalise and would therefore coincide with the result of [ki̯]. According to the traditional teleological perception of change, this would force an advance in the point of articulation of the result of [ki̯] to ensure its differentiation, which would have meant an approximation to the results of [ti̯] (Ariza, 2012, p. 55; Quilis, 2005, pp. 117–118). What seems clear, on the basis of the explanations consulted here, is that the evolution of the dental system is much more direct and transparent than that of the velar system, as can be seen in Figure 1.

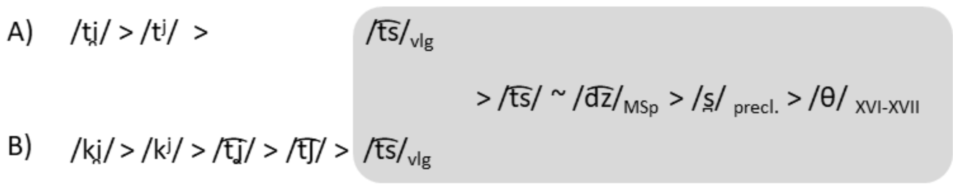


Figure 1. Diagram summarising the evolution of /ti̯/ and /ki̯/ yod contexts.

[2] This author indicates that the result of [ki̯] is a palatal affricate [t͡ʃ] (Zampaulo, 2019, p. 94): "The most frequent and direct result of [tj]-palatalisation is the affricate [ts], while that of [kj] most often renders [tʃ] (via *[cç])". On this point he agrees with Ariza (2012, pp. 28–29), who explains that the result of [ti̯] would originally have been [ʃ] and that of [ki̯], [c], results that would have yielded [t͡ʃ]. Subsequently, this affricate would have advanced its point of articulation to dentoalveolar to avoid a confluence with the evolution of /k, g/ followed by a palatal vowel.

However, there seem to be some differences as to when these sibilant pronunciations became voiced in intervocalic position (it is assumed that they remain voiceless after a consonant). Voicing is commonly considered to have taken place in mediaeval times (Lloyd, 1993, p. 427; Ariza, 2012, pp. 145–146); however, some authors point to a much earlier period: Lapesa (1981, p. 80) or Quilis (2005, p. 115) place this process in the Hispanic Latin of Imperial times. This would appear to be in line with the 4th century examples provided by Lloyd (1993) himself in describing the process of assibilation of the dental before a semivowel. In them, the spellings used in cases of confusion could suggest a certain degree of voicing: "VINCENTZA por VINCENTIA, LAURENTZIŌ por LAURENTIŌ" (Lloyd, 1993, p. 220).

Nevertheless, it does not seem likely that the shift from one phase to another is due to a discrete change; it is more logical to think that the substitution of one segment for another is due to the existence of allophonic variation, as Quilis (2005, pp. 115–116) has commented. The canonical exemplar at a given moment ends up being peripheral due to the previous existence of other variants, more so if there are intersections between exemplar clouds corresponding to different phonic categories (Pierrehumbert, 2001; Blevins, 2004; Recasens, 2011, p. 189). The processes of change leading to the appearance of the common assibilate realisations of [ti̯] and [ki̯] would thus have had a purely phonetic origin, based on articulatory, acoustic and, quite possibly, perceptual factors. In this sense, the approach of Recasens (2011) is very interesting: this author, after an articulatory analysis of sequences analogous to those of /ti̯, ki̯/ yod contexts for Catalan, concludes that the evolution has to do with coarticulatory processes that first involve the palatalisation of the stops and, subsequently, their affrication. This implies that modifications in the system take place unintentionally on the part of the speaker and that any supposed improvement or optimisation is, actually, a consequence of the speaker's action rather than a cause of it (Ohala, 1993, 2012; Blevins, 2004).

2.2. Palatalisation and assibilation processes

Yod processes are known to be associated with palatalisation. However, it is not always clear what palatalisation is, how it works and what it involves. The discussion remains active and notably intense (cf. Bhat, 1978; Halle, 2005, p. 23; Bateman, 2011; Repetti, 2016, p. 658). It has often been assumed that all processes involving a palatal element (a semivowel or a vowel), which influences and modifies another segment, fall under this term. This involves a vast array of phonetic processes ranging from coarticulatory assimilations to assibilation or rounding phenomena (Bhat, 1978). In this sense, the classification established by Bateman (2011, pp. 589–590), who distinguishes between full palatalisation and secondary palatalisation, is useful and precise. In the first case, a consonant adjacent to a palatal vowel or semivowel changes its point (and often its mode) of articulation, so that it occurs in the palatal region of the oral tract, i.e., a palatal consonant is obtained. In the second case, the consonant contiguous to a palatal vowel or semivowel segment acquires a palatal secondary articulation, but its primary point of articulation is not altered. The other phenomena traditionally associated with palatalisation could therefore not be considered as such, since the result would not be a segment with palatal characteristics. From this point of view, the assibilation affecting these yod contexts would not constitute a case of palatalisation in itself but would be the result of the previous process of secondary palatalisation.

This type of process is well studied in synchrony. Bhat (1978, p. 60) indicated that the presence of a palatal semivowel was one of the contexts most likely to induce palatalisation, especially in the case of apical consonants, such as dental consonants in Spanish. In these cases, the coarticulation of the consonant with the semivowel results in the partial assimilation of the palatal point of articulation by the consonant (Halle, 2005, p. 36; Recasens, 2011, pp. 189–193; Repetti, 2016, p. 658). In the present case, as Bateman (2011, p. 590) states, it would be more a matter of secondary palatalisation, as the original primary articulation is maintained. However, based on various studies (González Gelabert, 2019; Hall et al., 2006; Hall & Zygis, 2010, p. 18; Halle, 2005; Recasens, 2011), it has been shown that it is not uncommon for dental and velar stop consonants in these contexts to develop a certain degree of spirantization (assibilation), which leads to the appearance of assibilated elements, whether affricate or directly fricative. From this perspective, Hall et al. (2006, p. 60) define the assibilation of stops as "processes whereby stops become sibilant affricates or sibilant fricatives before high vocoids", which fits in perfectly with what has happened historically.

As Hall and Zygis (2010, pp. 18–19) point out, such phenomena should be seen as processes of consonantal lenition (cf. also Bybee & Easterday, 2019, p. 278), as they involve a reduction in the magnitude of articulatory gestures (the degree of constriction is less than in a stop): in front of palatal vowels and semivowels, assibilation (affrication) is more frequent due to the strong friction of the stop produced by the proximity of the tongue dorsum to the hard palate. The greater the coarticulation between the palatal segment and the stop, the greater the resulting friction and, consequently, the greater the assibilation.

It is important not to lose sight of what is meant by lenition, as this term is not without controversy. Here we apply the concept advanced by Bybee and Easterday (2019, p. 268), who relate it to phenomena involving the automation of articulatory gestures, typical in highly specialised motor patterns, which involve reduction and overlapping in the sequences of movements necessary for the production of sounds. These processes are what lead to efficiency in production: "the automatization process leads to systematic sound change, arising from coarticulation and reduction occurring in casual speech".

In other words, lenition is a reduction in the magnitude or duration of articulatory movements (Bybee & Easterday, 2019, p. 271), a stance compatible with the ideas of increased voicing and recourse to less marked forms alluded to by other scholars (cf. Kingston, 2008; Lavoie, 2001, p. 6). Thus, the transition to an affricate-like realisation in the history of Spanish would be justified as a weakening process resulting from the coarticulation of the stop consonant and the semivowel. In this sense, it should be remembered that lenition processes, according to Bybee and Easterday (2019, p. 288), are very frequent, are not restricted to a specific point of articulation and, in the specific case of affrication, one of the triggering contexts is the intervocalic position, specifically before a high and/or palatal segment.

There is evidence that the phonological contrast between coronal and dorsal consonants is neutralised before palatal semivowels, leading to confusion between velar and dental stops (Rallo & Fernández Planas, 1995; Halle, 2005, p. 37), which is consistent with what is postulated for diachronic change. Indeed, Chang et al. (2001, p. 100) conclude precisely that perceptual errors in sequences similar to those that caused the change exhibit the same asymmetries that had been observed diachronically: from [k] to [*t*][3] or, at least, to [t͡ʃ], but not the other way round. Such confusion would ultimately be caused by acoustic and perceptual factors.

2.3. Stop and sibilant consonants

Having reached this point, it is important to know how all these technical terms translate into the reality of sound, especially in the acoustic context, which is the perspective addressed here. As we know, the canonical Spanish stop consonants are characterised by the fact that they are produced in two successive phases: a first phase corresponding to the closure of the articulators (which results in a period of silence in the wave), and a release phase, corresponding to the sudden opening of the articulators forced by the air accumulated behind them (Quilis, 1993, pp. 206–207; Laver, 1994, p. 205;

[3] It should be noted that these authors evaluate English stop consonants, in which [*t*] is alveolar, not dental, hence the transcription without a diacritic.

Martínez Celdrán & Fernández Planas, 2007, pp. 31, 41–43). This phase translates acoustically into a release burst, whose characteristics can be good indicators of the point of articulation of the consonant (Martínez Celdrán & Fernández Planas, 2007, pp. 88–89). However, several works have documented allophonic variants that differ to varying degrees from this general portrait: it is possible to find stops without a clear release phase, i.e. only with an apparent closure phase and without a release burst, which can be classified as closed approximants (cf. Martínez Celdrán & Fernández Planas, 2007, p. 57; Martínez Celdrán, 2013, pp. 18–19). Stops with a certain degree of aspiration have also been documented, characterised by a slight friction at medium and low frequencies after the release burst (Ladefoged, 2003, p. 158; Ladefoged & Maddieson, 1996, pp. 66–70; O'Neill, 2010, pp. 24–26) or before it as preaspirated (Ladefoged & Maddieson, 1996, p. 70; Laver, 1994, p. 150; Parrell, 2012). Admittedly, the latter two variants have traditionally been associated with certain geographical varieties of Spanish, particularly southern ones (Alvar, 1961–1973, map 900; O'Neill, 2010; Parrell, 2012). Figure 2 gives an example of aspirated realisations in the southern variants, taken from O'Neill (2010, p. 24).

The study of the acoustic properties of stops reveals the existence of certain parameters that make it possible to differentiate their point of articulation. Traditionally, the onset frequency of F2 and F3 of the following sound, i.e. the starting point of the second and third transitions (T2 and T3, respectively, hereafter) is considered to be a good indicator of this feature, taking into account that T2 may vary depending on the following vowel. Specifically, in Spanish, Martínez Celdrán and Fernández Planas (2007, p. 91) place the onset frequency of T2 of [t̪] at just over 1700 Hz and that of [k] at just over 2000 Hz before the vowel [e], values that can serve as a guide if one takes into account that the following segment in the cases in question is also a palatal semivowel.[4]

Likewise, the duration of Voice Onset Time (VOT), i.e. "the distance from this [the burst] to the onset of laryngeal vibration" (Martínez Celdrán & Fernández Planas, 2007, p. 46), is recognised as a reliable index of the point of articulation: it lengthens as the point of articulation moves back, being around 20 ms for the voiceless dental and between 25 and 30 ms (or more) for the velar (cf. Martínez Celdrán & Fernández Planas, 2007, pp. 66–67, 89).

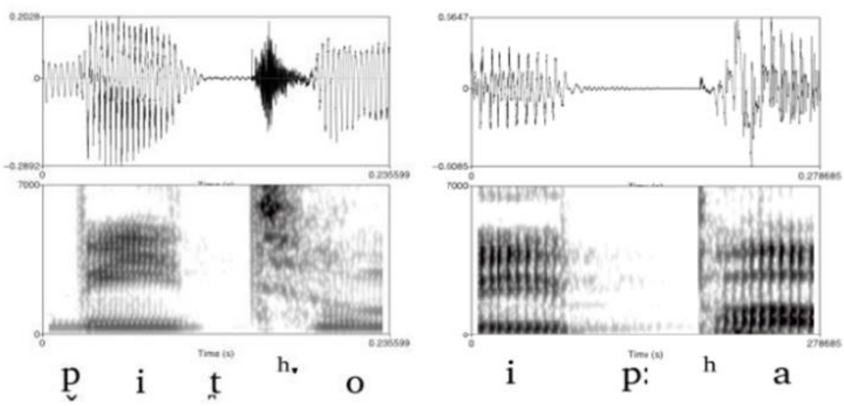


Figure 2. Examples of aspirated voiceless stops in Andalusian Spanish in the words *pisto* and *avispa* (O'Neill, 2010).

[4] With reference to the information provided by the transitions, it should be noted that much of the literature has focused on the analysis of the locus and/or locus equations (Ladefoged, 2003, p. 160; Martínez Celdrán & Fernández Planas, 2007, pp. 88–91; Quilis, 1993, pp. 210–211): there is agreement that after [t̪, d̪] T2 tends to converge towards 1700–1800 Hz; whereas, after [k, g], it would do so between 2000 and 3000 Hz (Harrington, 2010, p. 206; Quilis, 1993, p. 211). However, the usefulness of these parameters has been questioned (Ladefoged, 2003, pp. 160, 165; Martínez Celdrán & Fernández Planas, 2007, p. 92), so here we will take the initial frequencies of the two transitions into account.

The existence of assibilation can be detected and objectively examined from the duration of the opening phase, which is longer in cases where there is friction of some kind after separation of the articulators (Ladefoged & Maddieson, 1996, p. 90; Quilis, 1993, pp. 292–293).[5] Indeed, Hall et al. (2006, pp. 63–64) clearly explain that voiceless stops before palatal semivowels can have up to three distinct components: release burst (resulting from the abrupt opening of the articulators), local friction (resulting from turbulence caused by constriction in the supraglottic cavities) and aspiration (resulting from turbulence in the glottic cavity). The first appears just after the closure phase, the other two between the release burst and the next segment and they cannot always be easily identified. In general, friction shows a concentration of energy at high frequencies and can even be organised into something similar to formants. Aspiration, on the other hand, is usually noticed as faint energy in the lower frequencies (cf. also Ladefoged, 2003, p. 158; Ladefoged & Maddieson, 1996, pp. 66–70). These components may all appear or only some of them, resulting in the allophonic variation outlined above.

2.4. Obstruent voicing

Obstruent voicing in intervocalic contexts is one of the processes most commonly associated with consonant lenition (Bybee & Easterday, 2019; Lahoz, 2015, pp. 148–149; Lavoie, 2001, p. 30).

Bybee and Easterday (2019, p. 269) indicate that voicing is in fact one of the two main processes of weakening and define it as the process that "increases the vowel-like properties of a consonant, including both voicing and a decrease in the degree of constriction in the oral cavity".

In the present case, this type of phenomenon, often referred to as "passive voicing", is explained by a temporal misalignment of the start and end of vocal fold vibration (cf. Jessen & Ringen, 2002, p. 190; Westbury & Keating, 1986, p. 152).

Lahoz (2015) stated that a portion of the characteristic voicing of the surrounding sonorants can coincide with the voiceless segment, and any such overlapping could account for a significant proportion of its overall duration (p. 148).[6] This author maintains that, as a process of lenition, this would only be possible if there had previously been a reduction in consonant duration, which is quite common in the speech continuum of spontaneous communication situations (cf. García Santos, 2002, pp. 98–99; Martínez Celdrán & Fernández Planas, 2007, pp. 80–81). He also indicates that it is more likely in the middle position of an utterance than at its beginning or end, and in an unstressed context, since stressed syllables show a longer duration. Romero (1999) had already shown that Spanish voiceless stops tend to be voiced independently of their position in the word, at levels which, according to his explanation, would indicate that this type of segment is not actually totally voiceless in spontaneous speech and that voicing is not a categorical feature, but a continuous one, something that was also made clear in the work of Smith (1997) and Davidson (2016).

The same trends have been recognised for other languages. Arvaniti (2007, p. 106), building on earlier work by Nikolaidis, explains that, in Modern Greek stops, the classical descriptions that reduce the inventory to three voiced-voiceless pairs are not satisfactory, since the analysis of spontaneous speech reveals that processes of consonant lenition are at work, which, in the case of voiceless phonemes, take the form of voicing and, in both voiced and voiceless forms, a substantial variation in their point

[5] Quilis (199, pp.: 292–293) specifies that in fricative consonants the noise lasts about 110 ms; in affricates, 50 ms; and in aspirated occlusives, 30 ms. These values correspond to laboratory speech and are therefore hardly comparable with those of spontaneous, hypoarticulated, non-formal speech, which is characterised by a temporal reduction of segments (Lindblom, 1990).

[6] "parte de la sonorización características de las sonantes circundantes se puede solapar con el segmento sordo y […] cualquier solapamiento de este tipo puede suponer un porcentaje importante de su duración total" (Lahoz, 2015, pp. 148).

of articulation (as well as incomplete closure). This is fundamental to what we aim to investigate here, since it implies recognising that, in natural speech processes, phonological boundaries are blurred, and we see the emergence of a continuum which shows that voicing and alterations in mode and point of articulation go hand in hand.

3. Experimental design

In order to study the evolution of /ti̯/ and /ki̯/ yod contexts and to analyse the effects of the combination of consonant and palatal semivowel sequences, spontaneous speech recordings of between 45 minutes and one hour in length were used. These recordings, which form part of the GHECEM[7] corpus, correspond to free interviews on everyday topics carried out among peers, so the resulting degree of naturalness is very high. They were conducted in the Laboratory of the Department of Spanish, Modern and Classical Philology of the *Universitat de les Illes Balears* (UIB), with an Olympus LS-12 recorder placed about 50 cm from the respondents. Ten young speakers (6 men and 4 women), all of them undergraduate students at the UIB and fluent in Spanish, were interviewed.

All the cases of /ti̯e/ and /ki̯e/ were extracted from the speech samples obtained: the cases with syllable nuclei /a/ and /o/ were discarded as they were very rare and did not allow for a statistically reliable study[8]. The total number of examples included was 372 of /ti̯e/ and 160 of /ki̯e/, giving an overall total of 532 occurrences analysed, as shown in Table 1.

	/ti̯e/	**/ki̯e/**	**Total**
Speaker 1	46	13	59
Speaker 2	33	13	46
Speaker 3	20	15	35
Speaker 4	32	8	40
Speaker 5	28	24	52
Speaker 6	38	13	51
Speaker 7	23	15	38
Speaker 8	46	17	63
Speaker 9	50	16	66
Speaker 10	56	26	82
Total	372	160	532

Table 1. Number of cases analysed according to the type of sequence in each speaker.

The acoustic analysis of these sequences was carried out using oscillograms and broadband spectrograms with the *Praat* program (Boersma & Weenink, 1992–2021, v. 6.1.48), plus a script (Pérez Ramón, 2022) for the measurement of the parameters under study.

Different dependent variables were taken into account depending on the objective to be examined. To study the consonant assibilation process, the type of consonant allophone was taken into consideration, together with three quantitative parameters: the duration of the consonant opening phase and the onset frequencies of T2 and T3, since, as mentioned in §2.3, these are considered good indicators of the point of articulation of stops (Martínez Celdrán & Fernández Planas, 2007, pp. 88–91).[9] Two issues should be noted regarding these parameters: (a) in this work (following Hall et al., 2006) the duration of the entire opening phase has been measured, not

[7] *Grabaciones de Habla Espontánea en Catalán y Español de Mallorca* (*GHECEM*) is a corpus whose creation was undertaken in order to provide speech samples of sufficient quality to be analysed phonetically for the work of the FFI2014-59135-R project. Its initial aim is to study the way Catalan and Mallorcan Spanish influence each other. The result is a corpus which, at present, consists of 32 spontaneous speech samples of young Mallorcans with higher education, 9 of which are in Catalan (as a control) and the rest in Spanish (12 Spanish-dominant and 11 Catalan-dominant).

[8] For /ti̯/ followed by /a, o/ there were a total of 72 cases (17 with /a/ and 55 with /o/), while only 7 cases of /ki̯/ with /a, o/ were recorded. There are too few examples, and they are too unbalanced, to allow a solid comparison between contexts.

[9] It is important to bear in mind that the data referring to T2 and T3 also make it possible to identify the point of articulation of fricative sounds (cf. Quilis, 1993, p. 258; Marrero, in press; Martínez Celdrán & Fernández Planas, 2007, p. 108), so they will also be used in the case of allophony with segments of this type.

strictly the VOT, since the foreseeable presence of solutions with different degrees of voicing could skew the results, and (b) the frequency values have been normalised following the Lobanov method (Recasens, 2008) to neutralise possible effects of the sex of the speakers. As for the type of allophone, the criteria of Hall et al. (2006) have been applied in order to establish as objective a classification of them as possible, specified in six phonetic solutions according to their acoustic features: typical stops (closure with release burst), turbulent stops (closure with release burst and friction and/or aspiration), affricates (closure without release burst followed by friction) and sounds without a closure phase, which can be fricative, or open or closed approximants (in the case of having a formant structure)[10]. To study the voicing of the consonant, we adopted Smith's (1997, p. 478) criteria. Thus, the segments were classified according to their degree of devoicing into voiceless (between 100% and 76%), partially voiced (between 75% and 11%) and voiced (between 10% and 0%) realisations. The degree of devoicing has been determined with *Praat's Voice Report* function, which, among other parameters, provides the Fraction of Unvoiced Frames, which analyses the level of devoicing of a given sound.

The phonological type of sequence (/ti̯/ or /ki̯/) and the position in the word have been considered as factors, since an internal position tends to favour coarticulatory and lenitive processes in general, while prominent positions, such as the initial, tend to inhibit them (Lavoie, 2001, p. 27, also Westbury & Keating, 1986 for obstruent voicing). Thus, a distinction was made between initial position, post-consonantal internal position and intervocalic internal position. Stress has not been taken into account since all the cases identified are in the tonic syllable.

The statistical treatment of the data was carried out with SPSS (v. 25) using two types of model. Firstly, two multinomial logistic regression models with mixed effects were used, when the dependent variable was categorical (type of phonetic realisation and degree of voicing). The fixed effects were the type of sequence and the position in the word, with interaction between the two. Secondly, generalised linear models with mixed effects were used in cases where the dependent variable was quantitative (duration of the opening phase of the stop and T2 and T3 onset frequencies), also considering as fixed effects the type of sequence and the position in the word, as well as the interaction between the two. In all models, both the speaker and the word in which the sequence was found were treated as random variables. The significance value was set at 0.05.

4. Results

As previously mentioned, the results obtained pertain to two parallel major issues and are therefore presented in two sections: §4.1 presents the data associated with the process of assibilation of stop consonants (§4.1.1 presents the descriptive results and §4.1.2, the comparative results), while those relating to voicing are left for §4.2.

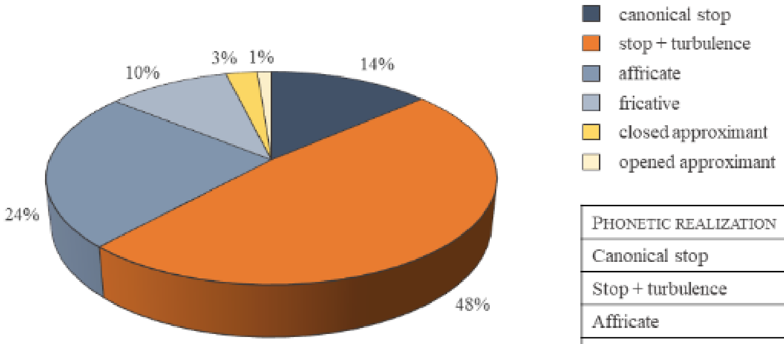
4.1. Assibilation

4.1.1. Descriptive results

4.1.1.1. Phonetic realisations of /ti̯/

In the case of the voiceless dental stop, 5 different phonetic realisations have been detected in addition to the typical stop [t̪], which was expected. The others correspond to stop consonants with turbulence, affricates, fricatives, open approximants and closed approximants. Interestingly, the commonest solution is by no means the prototypical dental stop (14.2% of the cases analysed), but rather the turbulent stop (48% of the total), followed by the affricate (24.4%). Far rarer are the cases of fricative consonants (9.9%), closed approximants (2.4%) and open approximants (1.1%). Hence, consonants with friction and/or aspiration are in the majority in this context, accounting for 82.3% of the total (Figure 3).

[10] The distinction between open and closed approximants established in Martínez Celdrán (2013), has been adhered to.



PHONETIC REALIZATION	utterances	relative frequency
Canonical stop	53	14.2%
Stop + turbulence	179	48%
Affricate	91	24.4%
Fricative	37	9.9%
Closed approximant	9	2.4%
Opened approximant	4	1.1%
TOTAL:	373	100%

Figure 3. Distribution of the allophones of /t/ in the sequence /ti̯/: the relative frequency is shown in the pie chart and the absolute frequency in the table.

Before proceeding, it is important to adequately describe each of these phonetic realisations. Detailed data are presented in Table 2 and Figure 4 (below). As can be seen in Figure 4.a, the two typical phases of this type of sound can be observed in the canonical stop: the initial silence phase followed by the release burst corresponding to the sudden opening of the articulators. This burst has an opening phase lasting 14.82 ms on average. The onset frequencies of T2 and T3 are 2263 Hz and 2910.71 Hz, respectively, higher than expected for a dental. Perceptually, they correspond to /t/. In contrast to typical dental stops, cases are found where, after the burst bar, clearly detectable in both the oscillogram and spectrogram, friction and/or aspiration appears, as described in Hall et al. (2006, pp. 63–64). In these, the opening phase of the sound is longer than in the canonical stop, with an average duration of 21.92 ms. It is not possible to refer directly to cases of aspiration, because this is not always the case: in some examples there is aspiration, while in others there is clear friction, or both occur simultaneously. In all cases, there is turbulence, as can be seen in Figure 4.b. The mean onset values of T2 and T3 are 2218.34 and 2876.83 Hz respectively. However, there are cases in which the appearance is that of an affricate consonant, in which, after the silent phase, friction (or aspiration) appears, without a visible release burst (Figure 4.c). The duration of the opening phase is equivalent to that of the previous allophone (22.11 ms on average); the onset frequency of T2 is 2077.13 Hz and that of T3 is 2782.47 Hz.

	Opening phase duration		**Initial frequency of T2**		**Initial frequency of T3**	
	x̄	***SD***	***x̄***	***SD***	***x̄***	***SD***
Stop	14.20	5.27	2263.00	273.44	2910.71	290.51
Stop with turbulence	21.92	7.76	2218.34	260.96	2876.83	251.19
Affricate	22.11	7.29	2077.13	291.04	2782.47	258.01
Fricative	—	—	1973.47	371.22	2781.09	289.46
Closed approximant	—	—	1942.89	438.18	2867.00	342.69
Open approximant	—	—	2104.25	55.75	2867.00	342.69

Table 2. Data for the opening phase duration (in ms) and T2 and T3 onset frequency (in Hz) for each of the realisations found for /ti̯/. The mean (x̄) and standard deviation (sd) are given.

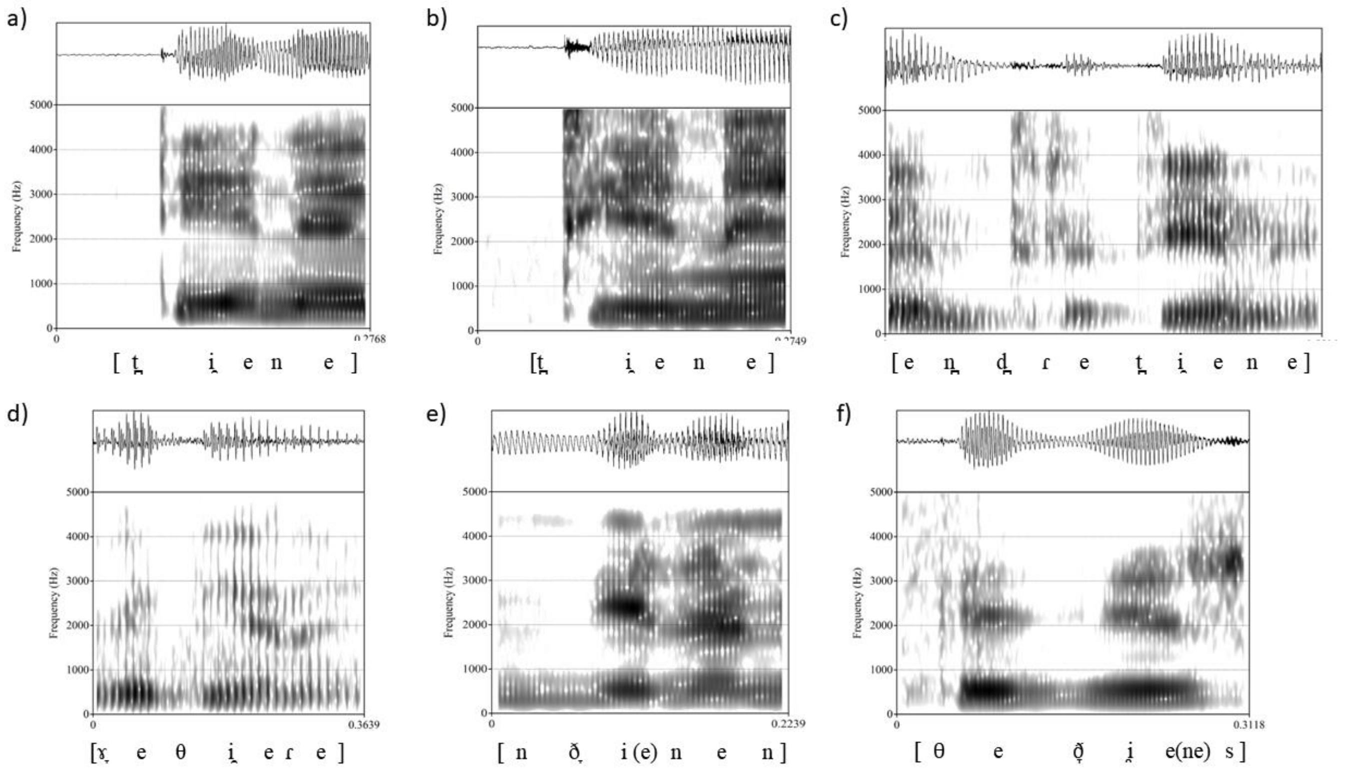


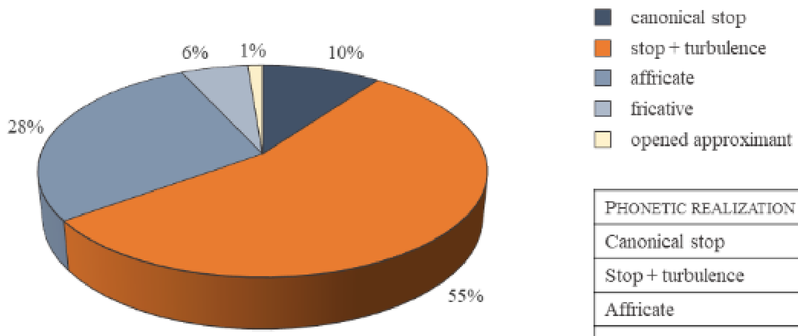
Figure 4. Charts with examples of the different allophones of /t/ in the sequence [ti̯]: (a) canonical stop (*tiene*), (b) turbulent stop (t*iene*), (c) affricate (*entretiene*), (d) fricative (*que tiene*), (e) closed approximant ([*tambié*]*n tienen*) and (f), open approximant (*te tienes*).

The last of the allophones that present turbulence are those cases in which there is no silent phase and only friction is seen (Figure 4.d), usually in the middle frequencies, something typical of fricatives in the interdental-dental range (Marrero, in press; Martínez Celdrán & Fernández Planas, 2007, pp. 106–110; Quilis, 1993, p. 264). Perceptually, they are clearly interpreted as fricatives of this type.

Finally, a few examples are given that cannot be classified as stops, fricatives or affricates because they have different acoustic features. Cases of closed approximants have been identified: as can be seen in Figure 4.e, in the oscillogram the wave is clearly periodic, there is voicing in the low frequencies, but formants are not observed, nor is there a release burst that could suggest a voiced stop. Furthermore, there are 4 cases we must consider as an open approximant (Figure 4.f): they show periodicity in the oscillogram, fairly distinct formants in the spectrogram and are clearly perceived as [ð̞].

4.1.1.2. Phonetic realisations of /ki̯/

Regarding the phonetic realisation of /ki̯/, as with the dental [ti̯], there is a considerable amount of variation documented. Again, the most frequent cases are not canonical velar stops (only 10% of the total), but rather examples involving friction of some kind: turbulent stops (55%), affricates (28.1%), and, much more rarely, fricatives (5.6%) and open approximants (1.3%), as can be seen in the data in Figure 5. Again, this implies that the most common acoustic manifestations preserve an occlusion prior to the opening of the articulators and that, in general, they usually present intermediate characteristics between a typical stop and an affricate. They are described in more detail below.



PHONETIC REALIZATION	utterances	Relative frequency
Canonical stop	16	10%
Stop + turbulence	88	55%
Affricate	45	28.1%
Fricative	9	5.6%
Opened approximant	2	1.3%
TOTAL:	160	100%

Figure 5. Distribution of the allophones of /k/ in the sequence /ki̯/: the relative frequency is shown in the pie chart and the absolute frequency in the table.

As expected, the cases of canonical stop respond to the classic description of this type of sound: they present a silent phase followed by a release burst, which often appears to be double (Figure 6.a). The opening phase is longer than that of the dental (18.19 ms on average), and the onset frequency of T2 and T3 corresponds to what is typical for a velar stop consonant (2212.8 Hz for T2 and 2810.87 Hz for T3). The complete data for the duration and onset frequency of T2 and T3 can be seen in Table 3 (below).

As in the case of the dentals, there are also examples where friction and/or aspiration can be detected after the easily noticeable release burst (Figure 6.b). In these cases, the duration of the opening phase is clearly longer than in dentals: 31.5 ms on average. The initial values of T2 and T3 are similar to those of the typical stop consonant: 2211.32 Hz and 2939.86 Hz, respectively. Cases are also found where, after silence, only friction or aspiration occurs, in a typical affricate configuration (Figure 6.c). The opening phase, as was already the case with the dentals, lasts practically the same length as in the previous type of allophones (31.28 ms) and the starting point of the transitions is also comparable (cf. Table 3). There are also a few cases in which the consonant typically corresponds to a fricative of the [x] type (Figure 6.d), as described in Marrero (in press), Martínez Celdrán and Fernández Planas (2007, pp. 106–108) and Quilis (1993, p. 270). Finally, two occurrences of open approximants were detected (Figure 6.e).

	Opening phase duration		**Initial frequency of T2**		**Initial frequency of T3**	
	$\bar{x}$	***SD***	$\bar{x}$	***SD***	$\bar{x}$	***SD***
Stop	18.19	10.42	2212.80	326.01	2810.87	210.55
Stop with turbulence	31.50	12.97	2211.32	451.80	2939.86	294.46
Affricate	31.28	11.27	2192.32	380.44	2974.11	249.32
Fricative	—	—	2217.33	207.74	2794.67	178.09
Open approximant	—	—	2115.00	340.82	2705.00	107.48

Table 3. Data for the opening phase duration data (in ms) and T2 and T3 onset frequency (in Hz) for each of the realisations found for /ki̯/. The mean ($\bar{x}$) and standard deviation (sd) are given.

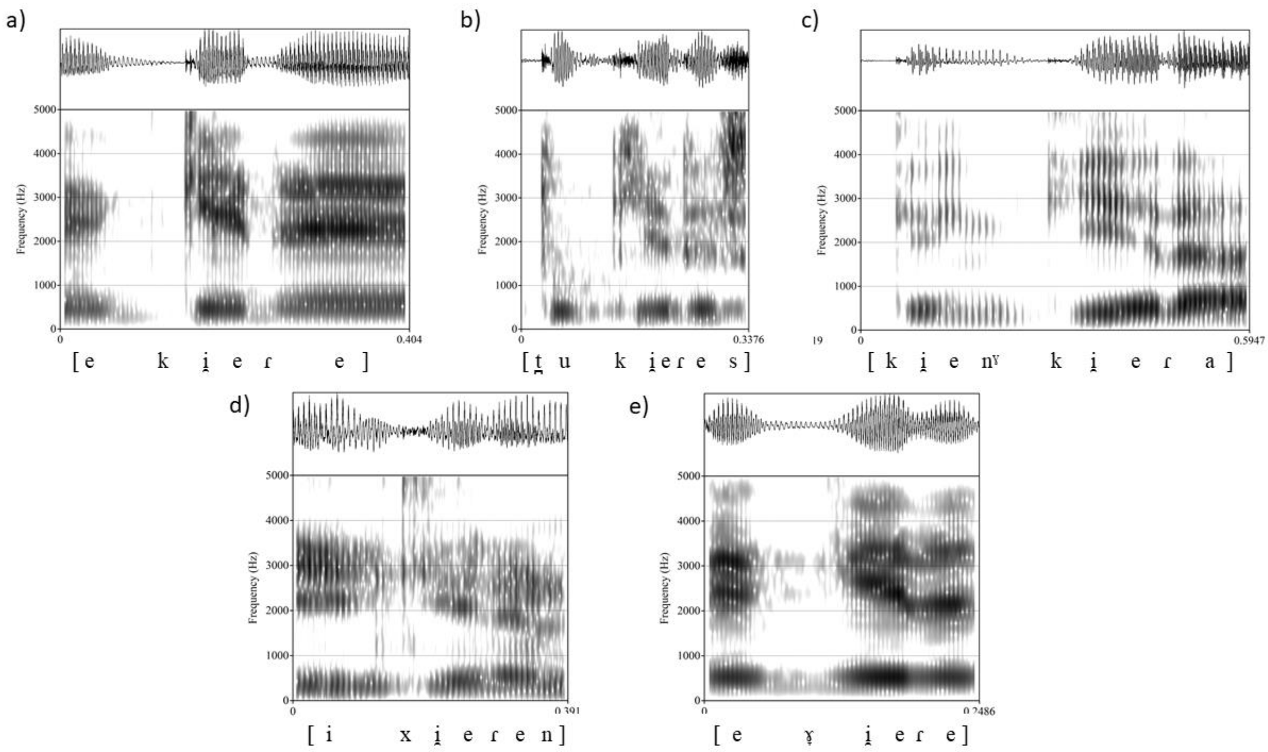


Figure 6. Examples of the allophones of /k/ in the sequence [ki̯]: (a) canonical stop ([*qu*]*e quiere*), (b) stop with turbulence (*tú quieres*), (c) affricate (*quienquiera*), (d) fricative ([*s*]*i quieren*), and (e) open approximant ([*qu*]*e* quiere).

4.1.2. Comparative results

Before proceeding, it is important to note that, in drawing up the inferential statistics, approximant realisations have not been considered, as there are too few examples to ensure reliable results.

The results of the statistical model show that the variants detected are conditioned neither by the type of sequence nor by their position in the word (see Table 4), which indicates that, from the point of view of phonetic solutions, there is a tendency towards convergence in the types of realisation and their distribution. However, the acoustic parameters of opening phase duration and the onset frequency of T2 and T3 need to be considered in order to determine more accurately whether this convergence of solutions is effective or apparent. Indeed, the statistics show that there are significant differences in some of these aspects.

The duration of the opening phase is related to the type of sequence [$F(1, 464) = 30.884$, $p < .0001$] (cf. Table 4). In general, it is significantly shorter for /ti̯/ than for /ki̯/ [$b_0 = 34.879$, $SE = 4.237$, $t = 8.232$, $p < .0001$], as would be expected from the phonetic characteristics of [t̪] and [k]: in the former case it is 10.82 ms shorter. A closer examination of the statistical data shows that, overall, the opening phase is shorter in initial than in internal post-consonantal position (3.43 ms less) [$SE = 1.743$, $t(464) = –1.970$, $p = .049$], and a certain degree of interaction between the type of sequence and the position in the word is detected, in that the acoustic manifestations of the dental consonant have a significantly shorter opening than those of the velar consonant, whatever its position: it lasts 9.26 ms less at the beginning of the word, 11.22 ms less in an interior position after the consonant and 12.70 ms less in an interior position between vowels. However, it should be noted that there is a speaker effect [*Est.* = 24.338, $SE = 12.135$, $Z = 2.006$, $p = .045$] and a word effect [*Est.* = 14.728, $SE = 6.014$, $Z = 2.449$, $p = .014$] that impact this behaviour.

Dependent variable	**Factors**		
	Sequence	**Position**	**Sequence × Position**
Allophone type	$F(4, 509) = 0.141$, $p = 0.967$	$F(8, 509) = 0.488$, $p = 0.865$	$F(8, 509) = 0.609$, $p = 0.771$
Duration	$F(1, 464) = 30.884$, $p < 0.0001$	$F(2, 464) = 2.504$, $p = 0.083$	$F(2, 464) = 0.348$, $p = 0.706$
Init. freq. T2	$F(1, 518) = 1.877$, $p = 0.171$	$F(2, 518) = 2.957$, $p = 0.053$	$F(2, 518) = 0.618$, $p = 0.539$
Init. freq. T3	$F(1, 518) = 1.431$, $p = 0.232$	$F(2, 518) = 0.523$, $p = 0.593$	$F(2, 518) = 0.198$, $p = 0.821$

Table 4. Results of multinomial logistic regression and generalised linear mixed-effects models for each of the dependent variables, and for their interaction.

Variables		**Allophone type**			
		Stop	**Stop + RB + Fric**	**Affricate**	**Fricative**
Sequence	**Duration**	$F(1, 62) = 4.789$, $p = 0.032$	$F(1, 261) = 46.292$, $p < 0.0001$	$F(1, 130) = 1.245$, $p = 0.267$	$F(4, 509) = 0.141$, $p = 0.967$
	Init. freq. T2	$F(1, 64) = 0.049$, $p = 0.825$	$F(1, 260) = 0.605$, $p = 0.437$	$F(1, 130) = 2.260$, $p = 0.135$	$F(1, 34) = 7.606$, $p = 0.009$
	Init. freq. T3	$F(1, 64) = 0.212$, $p = 0.647$	$F(1, 260) = 0.224$, $p = 0.636$	$F(1, 130) = 4.440$, $p = 0.037$	$F(1, 34) = 0.772$, $p = 0.386$
Position	**Duration**	$F(2, 62) = 6.807$, $p = 0.002$	$F(2, 261) = 3.297$, $p = 0.039$	$F(2, 130) = 1.215$, $p = 0.300$	$F(4, 509) = 0.141$, $p = 0.967$
	Init. freq. T2	$F(2, 64) = 1.289$, $p = 0.283$	$F(2, 260) = 1.199$, $p = 0.303$	$F(2, 130) = 0.542$, $p = 0.583$	$F(1, 34) = 1.902$, $p = 0.177$
	Init. freq. T3	$F(2, 64) = 0.080$, $p = 0.923$	$F(2, 260) = 1.013$, $p = 0.365$	$F(2, 130) = 1.493$, $p = 0.228$	$F(1, 34) = 0.245$, $p = 0.624$
Sequence × Position	**Duration**	$F(1, 62) = 2.203$, $p = 0.143$	$F(2, 261) = 3.258$, $p = 0.040$	$F(2, 130) = 4.751$, $p = 0.010$	$F(4, 509) = 0.141$, $p = 0.967$
	Init. freq. T2	$F(1, 64) = 0.003$, $p = 0.955$	$F(2, 260) = 0.859$, $p = 0.425$	$F(2, 130) = 0.478$, $p = 0.621$	$F(1, 34) = 1.977$, $p = 0.169$
	Init. freq. T3	$F(1, 64) = 1.211$, $p = 0.275$	$F(2, 260) = 0.450$, $p = 0.638$	$F(2, 130) = 0.037$, $p = 0.964$	—

Table 5. Results of the generalised linear mixed-effects models run for each type of allophone considered. Significant results are shaded.

With regard to the frequency of T2 and T3 onset, however, no clear link can be established either with the type of sequence or with the position (cf. Table 4). In both parameters, the frequency tends to be lower in initial position than in internal position, but this is only statistically significant in the case of T2 (0.306 points less in this position) [$SE = 0.137$, $t(518) = -2.229$, $p = .026$], although, in the case of T3, it is significant in the case of the dental consonant (0.298 points less than in the velar) [$SE = 0.095$, $t(518) = -3.132$, $p = .002$]. Again, a speaker effect is found in these results, both for T2 [$Est. = 0.218$, $SE = 0.111$, $Z = 1.970$, $p = .049$] and T3 [$Est. = 0.356$, $SE = 0.174$, $Z = 2.046$, $p = .041$].

These results indicate that, qualitatively, there is a match between the types of phonetic solution of /ti̯/ and those of /ki̯/, irrespective of their position. However, when analysing the articulation point indicator parameters, we find differences that allow us to distinguish phonetically between /ti̯/ and /ki̯/: in the first context, the opening phase would be shorter than in the second. Moreover, the onset frequency of T2 and T3 shows a certain relationship with the position in the word (lower in initial position), which can also be transferred to the type of sequence (in the case of T3, lower in the dental in that initial position). Therefore, there seem to be differentiated behaviours, although these are mostly influenced by the random effects of the speaker and the word in which the consonant is found. The existence of these divergences calls for a finer analysis, considering the type of allophone. The idea is to study equivalent allophones in /ti̯/ and /ki̯/ to see whether they are comparable (whether they show the same acoustic characteristics, regardless of their origin). The results of the statistical tests carried out for this purpose are given in Table 5 (above) and will be detailed in the following subsections. Figures 7 and 8 are a graphical synthesis of these results.

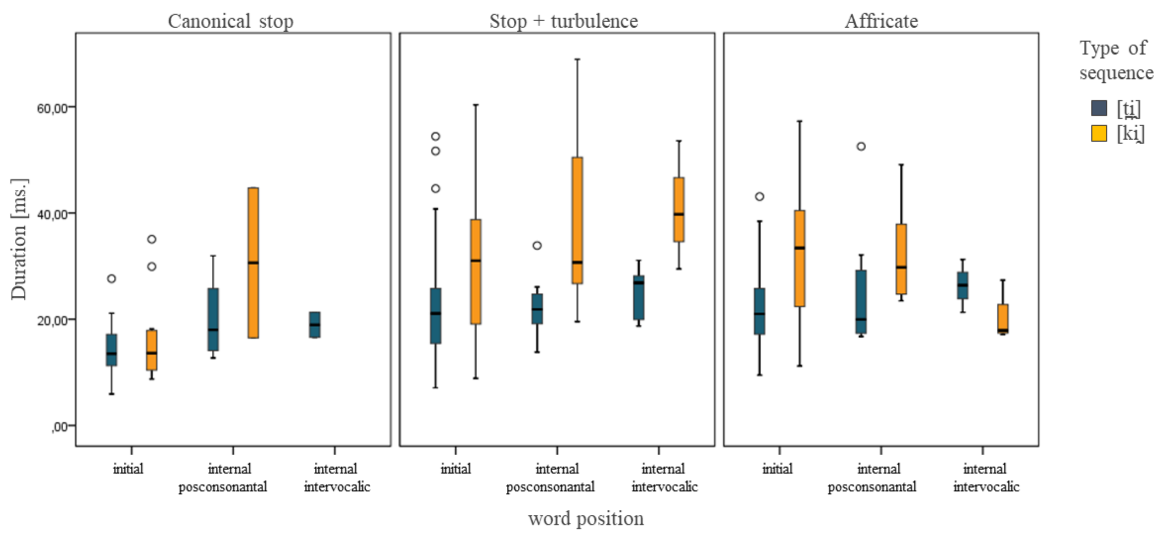


Figure 7. Box plot showing the differences in the duration of the opening phase of the different allophones, presenting it as a function of the type of sequence and the position in the word.

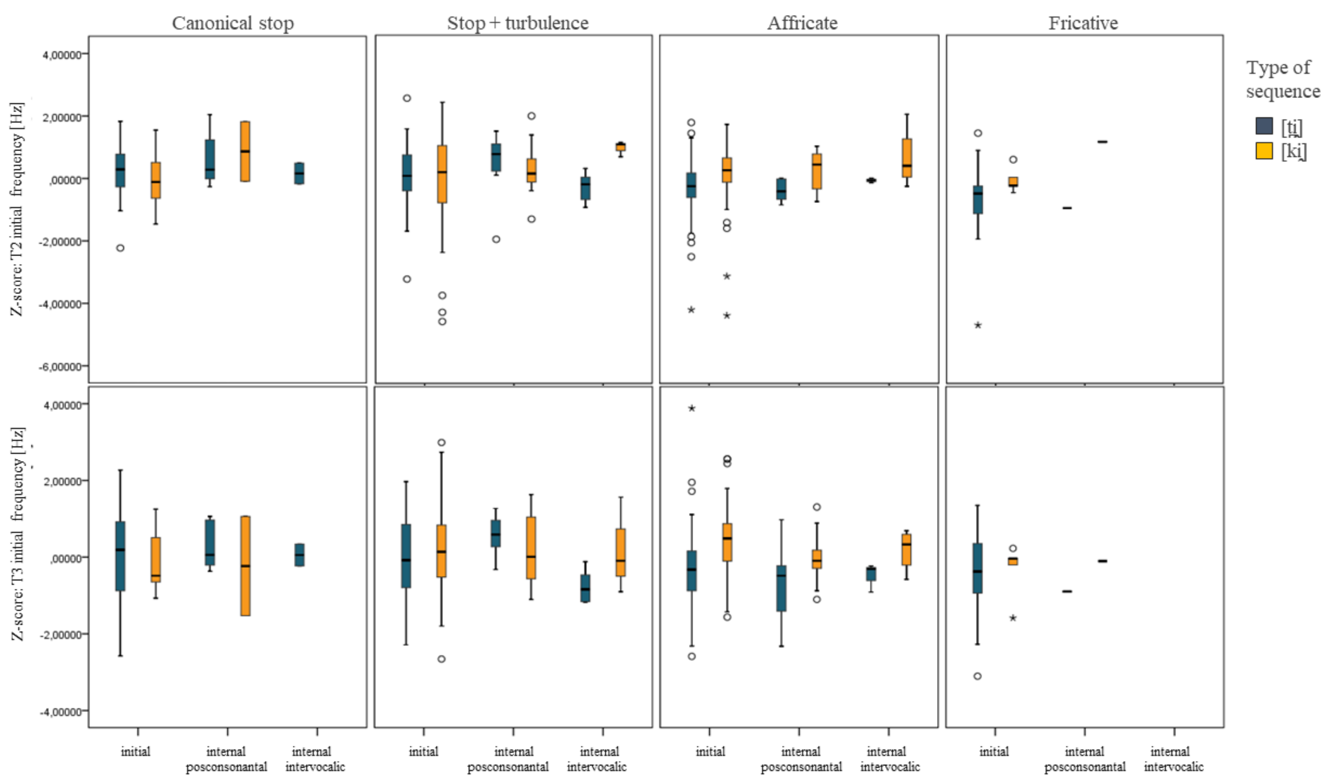


Figure 8. Box plots showing differences in T2 and T3 onset frequency (normalised data) as a function of sequence type and word position.

4.1.2.1. Typical stop consonant

The most prototypical stop consonants show significant differences in the duration of the opening phase, both for sequence type [$F(1, 62) = 4.789$, $p = .032$] and position [$F(2, 262) = 6.807$, $p = .002$]. In fact, the general trend outlined above is reproduced: in [k] the opening phase is significantly longer than in [t̪] [$b_0 = 28.509$, $SE = 6.374$, $t = 4.473$, $p < .0001$].

In addition, the opening is shorter in the initial position than in the internal post-consonant position, regardless of the sequence type (9.33 ms less) [$SE = 2.582$, $t(62) = –3.614$, $p = .001$] (see Figure 7). No interactions with the sequence have been detected, so this would be a general trend of the canonical stops examined here.

However, the frequency of T2 and T3 onset is not related to either sequence type or position, suggesting some overlap between [t̪] and [k] in this respect (see Figure 8).

4.1.2.2. Stops with turbulence

In the examples of stops with turbulence, statistically relevant differences are also detected relating the duration of the opening phase to the type of sequence [$F(1, 261) = 46.292$, $p < .0001$] and the position in the word [$F(2, 261) = 3.297$, $p = .039$], as was already the case for the more prototypical stops. The results indicate that the opening phase is significantly shorter for [t̪i̯] [$b_0 = –24.186$, $SE = 5.935$, $t = –4.075$, $p < .0001$] than for [ki̯] [$b_0 = 45.663$, $SE = 5.083$, $t = 8.983$, $p < .0001$]: there is a 15.72 ms difference between the two consonants. In general, it is significantly shorter in initial position (6.95 ms less than in intervocalic internal position) [$b_0 = –14.290$, $SE = 4.809$, $t = –2.972$, $p = .003$), something that was also observed in typical stops. Furthermore, a correlation between the two factors has been detected [$F(2, 261) = 3.258$, $p = .040$], which can be synthesised in the fact that the opening phase of [t̪i̯] has significantly shorter duration than that of [ki̯] in all the positions analysed: 9.50 ms less in initial position [$SE = 1.502$, $t(261) = –6.330$, $p < .0001$], 13.49 ms less in internal post-consonantal position [$SE = 3.373$, $t(261) = –3.999$, $p < .0001$], and 24.18 ms less in internal intervocalic position [$SE = 5.935$, $t(261) = –4.075$, $p < .0001$] (see Figure 7).

As in the case of the more canonical stops, these types of allophone do not show differences in the initial frequency of T2 and T3, either as a function of sequence type or of position, suggesting a coincidence in the point of articulation between [t̪] and [k]. It should be noted that there is a speaker effect on T3 onset frequency (*Est.* = 0.496, $SE = 0.244$, $Z = 2.035$, $p = .042$].

4.1.2.3. Affricates

As far as affricate-type sounds are concerned, the fixed effects (type of sequence and position) do not condition the duration of the opening phase, although there is a significant interaction between the two factors [$F(2, 130) = 4.751$, $p = .010$]. This initial statement must be qualified when examining the results in more detail. In general, it is observed that realisations of this type corresponding to /k/ have a significantly longer opening phase [$b_0 = 20.202$, $SE = 4.215$, $t = 4.793$, $p < 0.0001$). In initial ($b_0 = 13.074$, $SE = 4.312$, $t = 3.032$, $p = 0.003$) and internal post-consonantal ($b_0 = 10.820$, $SE = 4.769$, $t = 2.269$, $p = 0.025$) positions they also have a significantly longer duration. Interestingly, in this aspect the trend is the reverse of the previous two allophones, which points to a greater degree of assibilation (Figure 7). However, the aforementioned interaction reveals that the opening phase is significantly shorter in the case of initial [t̪i̯] ($b_0 = 20.202$, $SE = 4.215$, $t = 4.793$, $p < 0.0001$): 11.53 ms less compared to [ki̯] ($SE = 1.729$, $t(130) = –6.674$, $p < 0.0001$), which is consistent with documented trends in typical stops and stops with turbulence.

As far as the initial frequency of T2 is concerned, none of the factors taken into account have any influence, suggesting that the affricate solutions coincide in this aspect. However, the type of sequence does condition the frequency of T3 onset ($F(1, 130) = 4.440$, $p = 0.037$), significantly lower in the case of [ti̯] (0.66 points less; $SE = 0.316$, $t(130) = –2.107$, $p = 0.037$), which could point to a certain preservation of the differences in the point of articulation. Specifically, the difference with respect to [ki̯] is more prominent in the initial position (0.75 points less; *SE* = 0.222, *t*(130) = –3.399, *p* = 0.001).

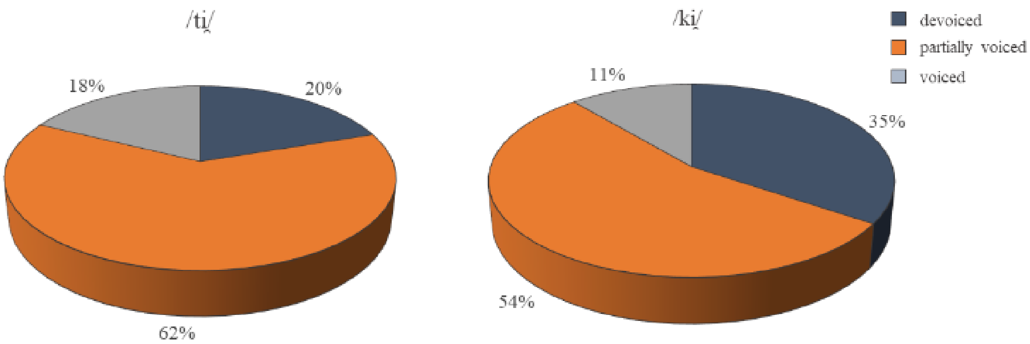
4.1.2.4. Fricatives

In the case of allophones consisting solely of friction, it was not possible to measure an opening phase as opposed to a closure phase of articulators, so only the frequency of onset of T2 and T3 was examined. The onset point of T2 is determined by the type of sequence ($F(1, 34) = 7.606$, $p = 0.009$): [t̪i̯] is significantly lower than [ki̯] (1.25 points less; b_0 = –1.873, *SE* = 0.843, *t* = –2.222, *p* < 0. 033), which is particularly noticeable in the initial (0.64 points lower; *SE* = 0.296, *t*(34) = –2.180, *p* = 0.036) and internal post-consonantal positions (1.87 points lower; *SE* = 0.843, *t*(34) = –2.222, *p* = 0.033), following a similar trend to the affricates (cf. Figure 8).

4.2. Obstruent voicing

Finally, we were interested in checking whether the voicing of the consonant was possible from the outset, taking into account, as the bibliography indicates, the position of the sequence in the word and the type of segment that precedes it as a triggering factor. Overall, it was observed that the sequences investigated can result in solutions with different degrees of voicing. Naturally, there are cases where the consonant remains completely voiceless (20% and 35% of the total in [t̪i̯] and [ki̯], respectively), but there are many cases of partial voicing (62% in [t̪i̯] and 54% in [ki̯]), and there are also a few cases where the consonant is fully voiced (18% in the dental and 11% in the velar), as can be seen in Figure 9.

The statistical model shows no difference in the degree of consonant voicing either as a function of sequence type or word position, which would seem to indicate that, in this respect, [t̪i̯] and [ki̯] behave similarly. However, it has been noted that the categorisation of the factor ‘word position’ is skewed with regard to the data collected for this experiment, since only 18 of the cases analysed correspond to an internal intervocalic position, while the majority correspond to word onset (459 of the total, 86.27%). It was therefore decided to readjust the analysis and consider the preceding segment but not the existence of a word boundary (except in cases where there is a preceding pause). Thus, three possible contexts have been established: after a pause, after a consonant and between vowels. The new statistical model reveals that the phonetic context determines the level of voicing of the consonant ($F(4, 526) = 13.123$, $p < 0.0001$): after a pause, solutions of the voiceless (b_0 = 2.658, *SE* = 0.692, *t* = 3.843, *p* < 0.0001) or partially voiced (b_0 = 1.927, *SE* = 0. 671, *t* = 2.872, *p* = 0.004) type are more likely than voiced ones; on the other hand, in intervocalic contexts, voiceless realisations are significantly less frequent than voiced ones (b_0 = –2.767, *SE* = 0.648, *t* = –4.270, *p* < 0.0001) (see Figure 10). Neither the speaker nor the word has any effect on voicing.



	devoiced	partially voiced	voiced	TOTAL
/t̪i̯/	74	233	66	373
/ki̯/	55	87	18	160

Figure 9. Distribution of the cases analysed according to their degree of voicing in the sequences under consideration.

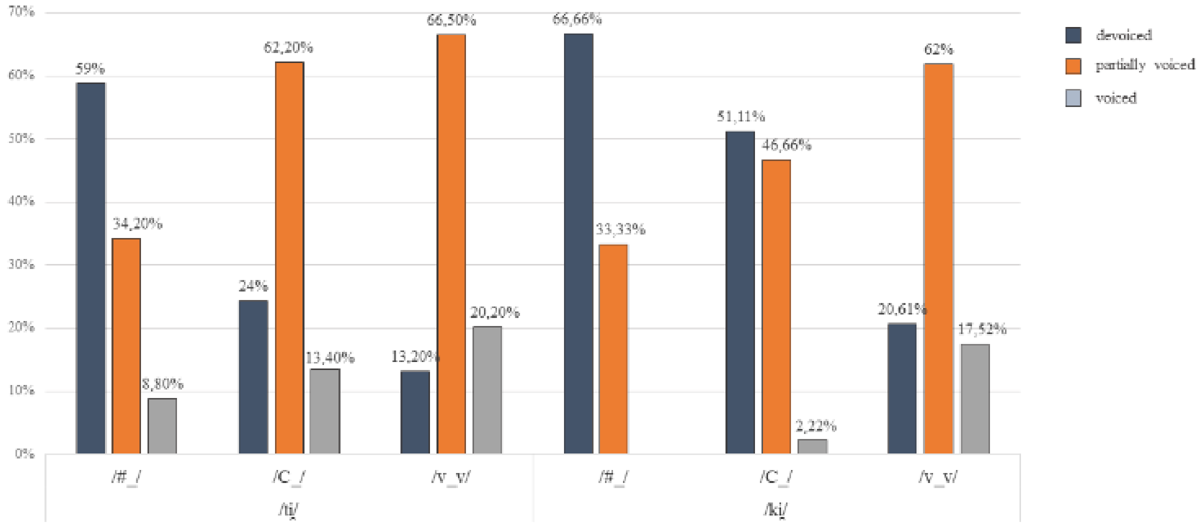


Figure 10. Bar chart showing the distribution of voiceless, partially voiced and voiced solutions as a function of phonetic context.

5. Discussion

The results obtained allow us to draw a very interesting picture both in terms of present-day Spanish and in terms of understanding the historical change of the yod contexts. The first thing to note, naturally, is that the data point to the existence of allophonic variation in the realisation of the stop consonant in the two types of sequence studied, as González Gelabert (2019) has already pointed out for /ti̯/.

What is significant is that the type of acoustic manifestation is very similar in both cases, encompassing solutions that range from the canonical sounds one would expect (typical stops with an explosion burst) to instances of relaxation that involve the emergence of approximants, encompassing segments with varying degrees of constriction and different levels of turbulence. It is true that /ti̯/ shows one more variant than /ki̯/, although this additional variant falls within the range of weakened solutions (the closed approximant) and constitutes a very small portion of the sample collected (1% of the total). We are thus dealing with solutions that range from more tense realisations with greater constriction (stops) to significantly weakened ones (approximants). This demonstrates a gradation from hyperarticulation to hypoarticulation in Lindblom's (1990) terms, indicating that lenition processes are also involved in this case and play a crucial role in the change (cf. García Santos, 2001). The variants detected can, therefore, be classified based on a progressive reduction in the magnitude and duration of articulatory gestures (following Bybee & Easterday's, 2019, definition of lenition, p. 271; cf. also Lavoie, 2001, pp. 20–21) as follows: typical stop > stop with turbulence > affricate > fricative > closed approximant > open approximant.

This situation shows that acoustic production responds to a phonetic continuum in which it is sometimes difficult to establish precise and clear boundaries between segmental categories (Recasens, 2011, p. 189). It is precisely for this reason, given the obvious similarities between the allophones found in both types of sequence, that we have chosen to classify them according to the acoustic elements that could be observed in oscillograms and spectrograms, following the model provided by Hall et al. (2006): the presence of closure, release burst, turbulence of some kind, and periodicity in the wave or formant structure are much more objective and clearer than a traditional cataloguing by mode and point of articulation, since these two parameters were not always easily identifiable.

However, the data reveal the existence of points of intersection in the production of the dental and velar consonants before the palatal semivowel. In short, significant variation is detected which can be organised, in Pierrehumbert's (2001, 2002) terms, into exemplar clouds arranged around a canonical or typical category (originally [t̪] for /ti̯/ and [k] for /ki̯/) according to their frequency of use. Some of these elements coincide at certain points with the exemplars of contiguous clouds in the acoustic-perceptual space. These points of coincidence, which had previously been highlighted in the literature (Chang et al., 2001; Hall et al., 2006; Recasens, 2011), are the result of the coarticulation of the two segments of the sequence, and they take the form of the presence of turbulence of a different nature (friction per se, aspiration or both at the same time, as already detected by Hall et al., 2006) in most of the cases analysed. Assibilation thus seems to be the rule rather than the exception, both in /ti̯/ and in /ki̯/. Following Pierrehumbert's (2001, 2002) explanations, the fact that these types of realisation are in the majority would mean that the central axis of these exemplar clouds would shift from the solutions considered canonical to those incorporating fricative features, which would then assume a central position. Thus, from a purely qualitative point of view, we find that the frequency of use of these much more common variants determines the change and would explain the progressive shift to [t͡s] in proto-Romance and, ultimately, to purely fricative solutions in the final phase of evolution. In this sense, it is very interesting to note that these fricative solutions are already present as possible variants of the stops in both /ti̯/ and /ki̯/: they are not, therefore, acoustic manifestations unrelated to the context or obligatory indicators of later phases in evolution. It is reasonable to presume that they would have existed from a very early stage, although their lower frequency of occurrence and their phonetic characteristics suggest much more relaxed phonemes (in which the closure phase has been lost), probably more common in very familiar communicative situations and, therefore, less associated with formal registers or styles. These would be clear cases, together with the examples of approximants, of hypoarticulation. Their prevalence in the final phase of evolution, however, is beyond the scope of this paper and needs further research.

These general data, therefore, come to corroborate, at least partially, what traditional historical phonetics defended (cf. Loporcaro, 2011, pp. 143–144; Repetti, 2016, p. 658; Zampaulo, 2019, pp. 94–95, among others): the acoustic nature of the variation detected confirms that the shift to a dentoalveolar affricate from a very early stage has a phonetic explanation based on the general aerodynamic mechanisms of production in the languages of the world, as reported by Hall et al. (2006, p. 61)[11]. However, they call into question the tendency to view the phases of change as hermetic and unvarying stages which follow one another in an orderly fashion as separate sections: the results show that the final solution already exists from the beginning and that the variation detected may condition the different phases due to questions of frequency of use that have to do with linguistic factors, internal to the system, and also with sociolinguistic factors (cf. Chambers, 2013, pp. 306, 316–317; Eckert & Labov, 2017; Pierrehumbert, 2002, p. 115).

Nevertheless, these considerations do not allow us to answer one of the questions posed here: whether the phonetic realisations of the stop consonants in /ti̯/ and /ki̯/ coincide beyond the features already mentioned; that is, whether we can speak of a real overlap in the production of the dental and the velar. For this purpose, three acoustic parameters were analysed: the duration of the opening phase and the onset frequencies of T2 and T3.

As we know, the duration of the opening phase allows us to address two questions: the process of assibilation itself, since it is a good indicator of the degree of affrication of the sounds, and the point of articulation, since it has been pointed out that the VOT (included in it) becomes longer as the point of articulation moves back (cf. Martínez Celdrán & Fernández Planas, 2007, pp. 66–67, 89). Regarding transitions, as has already been stated, it was

[11] "the creation of sibilants from stops has its phonetic origin in the brief period of turbulence (or 'friction phase') which occurs at the release of a stop into a following high vocoid" (Hall et al., 2006, p. 61).

decided to study their frequency of onset, following Martínez Celdrán & Fernández Planas (2007, p. 91). The general data obtained show that the duration of the opening phase allows us to distinguish between the dental and velar realisations in the sequences studied, since the former maintain a significantly shorter duration of the opening phase than the latter (20.73 ms on average vs. 30.05 ms), but not the onset frequencies of T2 and T3, which do not present statistically significant differences attributable to the point of articulation. In this regard, it is important to note that the onset point of T2 and T3 in the dental consonant is higher than expected for this type of sound according to the data of Martínez Celdrán and Fernández Planas (2007, p. 91): 2180.93 Hz for T2 and 2855.60 Hz for T3[12] (in fact, they coincide more with those provided by these authors for [k]). This, in principle, suggests that the phonetic realisation of the two stops is very similar in some aspects, specifically in these last two features, which are those most often related to the point of articulation. A certain degree of overlap between the two sequences can therefore be expected.

These data confirm in part what has been established by the diachronic literature about the initial results of the sequences /ti̯/ and /ki̯/: at first, different solutions would be found for each of these contexts (Ariza, 2012, p. 28; Lapesa, 1981, p. 79; Lloyd, 1993, p. 220); however, the existence of coincidences between them suggests the possibility of overlaps between their realisations, something that Quilis (2005, p. 116) has already defended. It is not unreasonable to think that these overlaps between the two types of sequence could lead to confusion and give rise to errors of interpretation, especially if we take into account that the few experiments that have tested perception of the point of articulation of voiceless stops have already shown the difficulties in the correct identification of /k/ and the likelihood of confusion with /t/ before a palatal vowel (cf. Chang et al., 2001; Rallo & Fernández Planas, 1995; Recasens, 2011, p. 217). Chang et al. (2001, p. 81), in fact, point out that the alteration of the acoustic features of the consonant implies problems in its categorisation.

The results show that these similarities occur directly from the two contexts equivalent to these yod clusters, without the need to consider concurrence with other similar sequence classes. Quite possibly, the allophonic variation resulting from cases of /t/ and /k/ followed by a palatal vowel may have contributed to this confusion, but it does not seem necessary to take them into account to explain the convergence in the results for /ti̯/ and /ki̯/.

The partial convergence between the phonetic realisation of /ti̯/ and that of /ki̯/, however, deserves a more detailed examination, considering the type of allophones detected. First of all, it can be observed that those sounds that exhibit closure with a release burst, i.e. a higher degree of constriction (typical stops and stops with turbulence), function in a similar way, in line with the general explanation offered above: their opening phase is significantly longer if they correspond to acoustic manifestations of /k/, which would allow their origin to be distinguished, but they do not present significant differences in terms of T2 and T3 onset frequency, which points to a certain overlap of features. In contrast, the affricate and fricative variants are less homogeneous and the differences that can be established between them are less systematic, suggesting that the lower the degree of constriction, the greater the degree of convergence between these solutions of /ti̯/ and /ki̯/. In the affricate, the opening phase by itself does not make it possible to distinguish between dental and velar, and only the onset frequency of T3 is a useful clue in this respect (2782.47 Hz in the case of dental vs. 2974.11 Hz in the case of velar). In the fricative, only the onset point of T2 makes it possible to differentiate between dental and velar (1973.47 Hz in the sequence /ti̯/, 2217.33 Hz in /ki̯/). The values obtained for the fricative corresponding to /t/, in fact,

[12] It is important to remember that the average obtained for the velar is 2174.13 Hz in T2 and 2946.28 Hz in T3. Nevertheless, as a perceptive anonymous reviewer rightly notes, the data in the literature are not directly comparable to those obtained here due to discrepancies in the experimental design (beginning with the type of participant); nonetheless, it has been deemed beneficial to utilize them as a starting point for an initial comparison.

are similar to some extent to those suggested in Carbó and Navarro (1999, p. 141) for [θ] before [e]; while those corresponding to the fricative obtained from /k/ coincide with those reported by Marrero (in press) for [x] [13].

There is another key element in the behaviour of the acoustic parameters analysed here which has been related to consonant lenition: position in the word. As already mentioned, syllable onset, and particularly word onset, is associated with a prominent position that tends to favour phonetic strengthening, while internal positions, especially intervocalic, are particularly prone to weakening (Recasens, 2011, pp. 213–214). Overall, it has been found that at word onset the opening phase tends to be significantly shorter than in internal positions, irrespective of the type of sequence. This is especially clear in the case of the dental consonant, the duration of whose opening phase is shorter in any position, with an ascending progression from initial position (20.54 ms on average) to post-consonantal internal (22.96 ms) and intervocalic internal (24.16 ms). Regarding the velar consonant, this gradation varies: the opening is shorter in the initial position (28.96 ms), and progressively longer in the intervocalic internal (29 ms) and post-consonantal internal (34.51 ms). Nonetheless, this implies that, in line with expectations, an initial position would inhibit relaxation in consonant articulation and hinder assibilation, which usually requires a longer opening phase for turbulence to occur. On the other hand, in an internal and less prosodically prominent position, lenition is more feasible and so, therefore, assibilation, which entails an increase in the duration of the opening phase, would be more likely. The same behaviour is detected in the case of the allophones classified as typical stops and in stops with turbulence, as well as in the affricate realisation of /ti̯/. Curiously, in /ki̯/ the tendency is the opposite: its opening is comparatively longer in initial and post-consonantal position. In both affricates and fricatives, there are differences in the frequency of onset of transitions depending on this variable: in the former, T3 is lower in initial position, while in the latter, T2 is lower.

The results outlined so far allow us to understand the process of assibilation and to make initial conjectures about the convergence between the solutions of these two yod contexts. Nevertheless, voicing of these sequences has yet to be discussed. The data collected show that most of the phonetic solutions produced are not strictly voiceless, but partially voiced. If we take into account that there are sound variants according to the scale of voicing established by Smith (1997, p. 478), the examples in which voicing can be detected in proportions that are far from what is typical of unvoiced sounds are in the vast majority. The tendency is equally unequivocal for both /ti̯/ and /ki̯/. This suggests that voicing may not have followed assibilation, but that both phenomena occurred simultaneously, as Arvaniti (2007, p. 106) commented for Modern Greek. It is important to remember that both assibilation and voicing are, ultimately, processes of consonant lenition (one by reduction in articulatory constriction in the oral cavity, the other by reduction in duration and consequent temporal misalignment), as authors such as Bybee and Easterday (2019, pp. 270–272), García Santos (2002) and Lavoie (2001, pp. 20–21) have pointed out. In this regard, the type of adjacent segment has a clear influence on the weakening process: it has been shown that a post-pausal context favours the emergence of voiceless or partially voiced solutions, while the intervocalic context inhibits them in favour of voiced ones, which are more frequent in this context. Lahoz (2015, pp. 148–150) argued that this passive voicing is conditioned by the temporal reduction of the consonant, an indispensable precondition for the voicing of the adjoining segments to be misaligned and transferred to the consonant in question. This is highly relevant to the cases examined here, since we have used samples of spontaneous speech (the condition in which sound changes occur), one of the

[13] The author refers the reader to the data of Borzone de Manrique and Massone (1981), who place the T2 frequency of [x] at 2180 Hz, similar to that obtained here.

most commonly emphasised features of which is the shorter duration of segments, a feature of the hypo-articulation inherent in this type of discourse (Lindblom, 1990). It should not be overlooked that the recordings were made among peers (there were relationships of friendship or kinship), the conversation dealt with topics of common and everyday interest for both interviewer and interviewee, so the degree of trust and familiarity between the two was considerable, resulting in the collection of very natural speech samples.

Despite all that has been said, this paper is only a first approximation to this question. From the point of view of acoustic analysis, much remains to be verified, especially in terms of more accurately determining the point of articulation of the realisations documented. It is therefore necessary to examine spectral moments in the cases where turbulence is observed, to examine the data for intensity and the spectral peaks in the release burst of the segments that show turbulence, and to study the transitions in more detail. In view of the variation in the examples collected, it is also relevant to address the role that the simplification of the diphthong may have played in the acoustic characteristics of the consonant. Furthermore, a fundamental part of understanding the final convergence between /ti̯/ and /ki̯/ remains to be addressed: its auditory discrimination. In this regard, it is essential to carry out perceptual tests to determine whether perception, as is foreseeable given the dynamics of the processes of sound change (Blevins, 2004; Ohala, 2012; Pierrehumbert, 2001, 2002), played a key role in their development. This would also be particularly interesting for the last stage of the historical process in which the fricative variants finally prevailed, having been part of the exemplar cloud from the beginning.

6. Conclusions

This study is only an initial phase in the explanation of the diachronic evolution of the /ti̯, ki̯/ yod clusters from the perspective of acoustic phonetics. Traditionally, it had been accounted for from the articulatory perspective, but the acoustic properties of the sounds and their perceptual interpretation had been overlooked. The analyses carried out allow some preliminary conclusions to be drawn regarding this process of change, but also make it possible to refocus the idea of change as a progressive succession of discrete stages. It is evident from the results that the realisation of the sequences that trigger the evolution of /ti̯/ and /ki̯/ are not phonetically uniform, but show a very important allophonic variation that can be organised in a continuum from greater to lesser tension (typical stop segments > turbulent stops > affricates > fricatives > approximants). This implies that evolution responds to patterns of reduction in the magnitude of articulatory gestures, along the lines described by Bybee and Easterday (2019), Kingston (2008) and Lavoie (2001) and supported by García Santos (2001, 2002).

Among these phonetic solutions for /t/ and for /k/, it is striking that we find examples of all the stages through which it has been established that change occurs over the centuries. This suggests that rather than new phonetic realisations occurring over time, what really changes is their distribution in use. Following Pierrehumbert (2001), what at first were marginal solutions within the group of exemplars would have shifted to a central position in that cloud and would have ended up being considered canonical, to the detriment of others. The selection of one or the other as prototypical variants would have been influenced, initially and based on the result of evolution, by hypoarticulation issues, although extra-linguistic factors must necessarily have played a role in their diffusion. It has, moreover, been confirmed that the acoustic manifestations derived from /t/ and /k/ before the palatal semivowel were originally partially different, from which we may conclude that the segments are not totally different but have enough similarities to end up converging, possibly for perceptual reasons.

Finally, the preliminary analysis of voicing carried out here raises the need to reconsider the process of voicing in these segments, which had been considered as typical of mediaeval Castilian Spanish. The results obtained show that the voicing may have occurred from a very early stage, favoured by intervocalic phonetic contexts and, possibly, due to a previous reduction in the duration of the consonant typical of a spontaneous speech situation, which is where the change takes place.

Much remains to be considered and studied, but this initial approach makes it possible to lay the foundations for more solid research and to raise some questions whose resolution should lead to a better understanding of a process of phonic change which, although it has been widely studied, still justifies further analysis.

Acknowledgements

This study has been financed by the Comunitat Autonòma de les Illes Balears through the Direcció General de Política Universitària i Recerca with funds from the Ley de Estancias Turísticas (PDR2020/51 - ITS2017-006). It also benefits from project TED2021-2021-130752A-I00. This work would have been much more challenging without the voluntary help of two students from the Universitat de les Illes Balears, Cati Riber and Isabel Sancho de la Jordana, who carried out the segmentation of the sound waves.

References

Alvar, M. (1961–1973). *Atlas lingüístico y etnográfico de Andalucía.* Universidad de Granada/CSIC. (In cooperation with A. Llorente and G. Salvador.)

Ariza, M. (2012). *Fonología y fonética históricas del español*. Arco Libros.

Arvaniti, A. (2007). Greek Phonetics: The state of the art. *Journal of Greek Linguistics*, *8*, 97–208. https://doi.org/10.1075/jgl.8.08arv

Bateman, N. (2011). On the typology of palatalization. *Language and Linguistic Compass*, *5*(8), 588–602. https://doi.org/10.1111/j.1749-818X.2011.00294.x

Bhat, D. N. S. (1978). A general study of palatalization. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of Human Language* (vol. 2, pp. 47–92). Stanford University Press.

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486357

Boersma, P., & Weenink, D. (1992–2021). *Praat: Doing phonetics by computer* (Version 6.1.48) [Computer program]. https://www.praat.org/

Borzone de Manrique, A. M., & Massone, M. I. (1981). Acoustic analysis and perception of Spanish fricative consonants. *Journal of the Acoustical Society of America*, *69*(4), 1145–1153. https://doi.org/10.1121/1.385694

Bybee, J., & Easterday, S. (2019). Consonant strengthening: A crosslinguistic survey and articulatory proposal. *Linguistic Typology*, *23*(2), 263–302. https://doi.org/10.1515/lingty-2019-0015

Carbó, C., & Navarro, M. J. (1999). Características acústicas de los sonidos fricativos del español. In E. Martínez Celdrán, S. Planas Morales, & J. Romero Gallego (Eds.), *Actes del I Congrés de Fonètica Experimental* (pp. 135–142). Universitat Rovira i Virgili – Universitat de Barcelona.

Chambers, J. K. (2013). Patterns of variation including change. In J. K. Chambers, & N. Schilling (Eds.), *The Handbook of Language Variation and Change* (pp. 297–323). Wiley. https://doi.org/10.1002/9781118335598.ch14

Chang, S. S., Plauché M. C., & Ohala, J. J. (2001). Markedness and consonant confusion asymmetries. In E. Hume, & K. Johnson (Eds.), *The Role of Speech Perception in Phonology* (pp. 79–101). Academic Press.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, *54*, 35–50. https://doi.org/10.1016/j.wocn.2015.09.003

De Vogelaer, G., & Seiler, G. (2012). The dialect laboratory: Introductory remarks. In G. De Vogelaer, & G. Seiler (Eds.), *The Dialect Laboratory* (pp. 1–32). John Benjamins.

Eckert, P., & Labov, W. (2017). Phonetic, phonology and social meaning. *Journal of Sociolinguistics*, *21*(4), 467–496. https://doi.org/10.1111/josl.12244

García Santos, J. F. (2001). Ni velarización, ni interdentalización: Lenición. In J. Bartol

Hernández (Ed.), *Nuevas aportaciones al estudio de la lengua española: Investigaciones filológicas* (pp. 93–102). Luso-Española Ediciones.

García Santos, J. F. (2002). *Cambio fonético y fonética acústica*. Universidad de Salamanca.

González Gelabert, R. (2019). El cambio fónico de yod primera: La evolución histórica a través del español actual. *Res Diachronicae*, *17*, 22–45.

Guion, S. G. (1998). The role of perception and velar palatalization. *Phonetica*, *55*(1–2), 18–52. https://doi.org/10.1159/000028423

Hall, T. A., & Zygis, M. (2010). An overview of the phonology of obstruents. In S. Fuchs, M. Toda, & M. Zygis (Eds.), *Turbulent sounds: An interdisciplinary guide* (pp. 1–36). De Guyter. https://doi.org/10.1515/9783110226584.1

Hall, T. A., Hamann, S., & Zygis, M. (2006). The phonetic motivation for phonological stop assibilation. *Journal of the International Phonetic Association*, *36*(1), 59–81. https://doi.org/10.1017/S0025100306002453

Halle, M. (2005). Palatalization/velar softening: What is and what it tells us about the nature of language. *Linguistic Inquiry*, *36*(1), 23–41. https://doi.org/10.1162/0024389052993673

Harrington, J. (2010). *Phonetic Analysis of Speech Corpora*. Wiley-Blackwell.

Harrington, J. (2012). The relationship between synchronic variation and diachronic change: Modeling exemplar-based phonologization. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 321–332). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199575039.013.0013

Hinskens, F. (2021). The expanding universe of the study of sound change. In R. D. Janda, B. D. Joseph, & B. S. Vance (Eds.), *The Handbook of Historical Linguistics* (vol. 2, pp. 7–46). Wiley. https://doi.org/10.1002/9781118732168.ch2

Jessen, M., & Ringen, C. (2002). Laryngeal features in German. *Phonology*, *19*, 189–218. https://doi.org/10.1017/S0952675702004311

Kingston, J. (2008). Lenition. In L. Colantoni, & J. Steele (Eds.), *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 1–31). Cascadilla Press.

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Blackwell.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.

Lahoz, J. M. (2015). *Fonética y fonología de los fenómenos de refuerzo consonántico en el seno de unidades léxicas en español*. [Doctoral dissertation, Universidad Complutense de Madrid – Instituto Universitario de Investigación Ortega y Gasset]. Docta Complutense. https://hdl.handle.net/20.500.14352/38657

Lapesa, R. (1981). *Historia de la lengua española*. Gredos.

Laver, J. (1994). *Principles of Phonetics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139166621

Lavoie, L. (2001). *Consonant Strength: Phonological Patterns and Phonetic Manifestations*. Routledge. https://doi.org/10.4324/9780203826423

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403–439). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-009-2037-8_16

Lloyd, P. M. (1993). *Del latín al español: I. Fonología y morfología históricas de la lengua española*. Gredos.

Loporcaro, M. (2011). Phonological processes. In M. Maiden, J. C. Smith, & A. Ledgeway (Eds.), *The Cambridge History of the Romance Languages* (vol. 1, pp. 109–154). Cambridge University Press. https://doi.org/10.1017/CHOL9780521800723.005

Marrero, V. (En prensa). Descripción fonética de las fricativas y las africadas. In J. Gil, & J. Llisterri (Eds.), *Fonética y fonología descriptivas de la lengua española*. Georgetown University Press.

Martínez Celdrán, E. (2013). Caracterización acústica de las aproximantes espirantes en español. *Estudios de Fonética Experimental*, *22*, 11–35. https://raco.cat/index.php/EFE/article/view/275406

Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española: Articulaciones y sonidos del español*. Ariel.

Menéndez Pidal, R. (1904/1985). *Manual de gramática histórica española* (18th ed.). Espasa-Calpe.

O'Neill, P. (2010). Variación y cambio en las consonantes oclusivas del español de Andalucía. *Estudios de Fonética Experimental*, *19*, 11–41. https://raco.cat/index.php/EFE/article/view/218540

Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical Linguistics: Problems and Perspectives* (pp. 237–278). Longman.

Ohala, J. J. (2012). The listener as a source of sound change: An update. In M.-J. Solé, & D. Recasens (Eds.), *The Initiation of Sound Change* (pp. 21–35). John Benjamins. https://doi.org/10.1075/cilt.323.05oha

Parrell, B. (2012). The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of Phonetics*, *40*(1), 37–45. https://doi.org/10.1016/j.wocn.2011.08.004

Pérez Ramón, R. (2022). Script para datos de duración VOT, frecuencia inicial de T2 y T3. [Unpublished Praat script].

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee, & P. Hopper (Eds.), *Frequency Effects and the Emergence of Linguistic Structure* (pp. 137–151). John Benjamins. https://psycnet.apa.org/doi/10.1075/tsl.45.08pie

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven, & N. Warner (Eds.). *Laboratory Phonology 7* (pp. 101–139). De Gruyter. https://doi.org/10.1515/9783110197105.1.101

Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Gredos.

Quilis, A. (2005). *Fonética histórica y fonología diacrónica*. UNED.

Rallo, L., & Fernández Planas, A. M. (1995). La invariación acústica en las oclusivas sordas del castellano: Estudio perceptivo. *Estudios de Fonética Experimental*, *7*, 45–74. https://raco.cat/index.php/EFE/article/view/144408

Recasens, D. (2008). Mètodes de normalització i de representació de dades acústiques i articulatòries. *Estudios de Fonética Experimental*, *17*, 331–341. https://raco.cat/index.php/EFE/article/view/140076

Recasens, D. (2011). Velar and dental stop consonant softening in Romance. *Diachronica*, *28*(2), 186–224. https://doi.org/10.1075/dia.28.2.02rec

Repetti, L. (2016). Palatalization. In A. Ledgeway, & M. Maiden (Eds.), *The Oxford Guide to the Romance Languages* (pp. 658–668). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199677108.003.0039

Romero, J. (1999). The effect of voicing assimilation on gestural coordination. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1793–1796).

Smith, C. L. (1997). The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics*, *25*(4), 471–500. https://doi.org/10.1006/jpho.1997.0053

Westbury, J. R., & Keating, P. (1986). On the naturalness of stop consonants voicing. *Journal of Linguistics*, *22*(1), 145–166. https://doi.org/10.1017/S0022226700010598

Zampaulo, A. (2019). *Palatal Sound Change in the Romance Languages. Diachronic and Synchronic Perspectives*. Oxford University Press. https://doi.org/10.1093/oso/9780198807384.001.0001