

# Ibn Sīnā's Remarks on a Khwarizmian Sound\*



ADAM BENKATO 

## Abstract

In his study of Arabic phonetics, *Asbāb ḥudūt al-ḥurūf* (*The Causes of the Genesis of the Consonants*), Ibn Sīnā briefly surveys some speech sounds found in languages other than Arabic, among them one particular to Khwarizmian, an Iranian language attested primarily in glosses to Arabic manuscripts of the 13th century. This study attempts to elucidate the sound Ibn Sīnā describes both through reference to his own system of phonetic terminology and through comparison with extant material in the Khwarizmian language.

**Keywords:** Khwarizmian; Middle Iranian; Arabic; Avicenna; Al-Biruni

The Khwarizmian language, belonging to the Iranian branch of the Indo-European family and spoken in the fertile delta of the Amu Darya river south of the Aral Sea, was long known to have existed only through the reports of the famed polymath Abū Rayḥān al-Bīrūnī (d. 1050). In one of his most important extant works, *al-Aṭār al-bāqīya* ‘*an al-qurūn al-khālīya* (Chronology), he discusses various calendrical terms, giving the names of the months, days, and lunar stations in Khwarizmian as he does for Arabic, Hebrew, Persian, Syriac, and Sogdian.<sup>1</sup> In the same work, al-Bīrūnī also laments the Arab conquest of Khwarizm which led to the destruction of older institutions, especially to a loss of the knowledge of writing in

\*It is a pleasure to offer this study to François de Blois, with whom I studied Khwarizmian some years ago, in honour of his incomparable scholarship at the intersections of Arabic and Iranian philology. For drawing my attention to the remarks of Ibn Sīnā discussed herein and commenting helpfully on a draft of this paper, I thank Kevin Van Bladel.

<sup>1</sup>Sachau, Eduard, *The Chronology of Ancient Nations* (London, 1879), pp. 57–58 and 223–228. The name of the region and the language have been rendered variously in works in European languages: “Chorasmian” and “Chorasmian” are based on Greek Χορασμία while “Khwarezmian” is based on Persian خوارزم and “Khwarazmian” seems to be based on the Old Persian form *(h)wārazmiš*. The Arabic form occurring in the extant textual sources is *khwarizm* خوارزم. For consistency with the Arabic-language source material, I use Khwarizm/Khwarizmian throughout this piece. The name of the region itself is attested as Avestan as *xwārizəm* (acc.sg.) in addition to Old Persian as *(h)wārazmiš* <u-v-a-r-z-m<sup>1</sup>-i-š> (nom.sg.), which may be derived from \*hwāra- “low” (as was already recognized by David MacKenzie (‘Khwarazmian language and literature’, in *The Cambridge History of Iran* Vol. III, Part 2 (1983), p. 1244) plus \*zm- (zero-grade of \*zam- “land”) with a likely meaning of “low-lands”—not unreasonable given the low elevation of the marshy (in antiquity) region south of the Aral Sea.

the indigenous Khwarizmian script.<sup>2</sup> As a Khwarizmian by birth, al-Bīrūnī was of course personally familiar with its history and customs.<sup>3</sup> On the basis of a passage in another of his works, the *Kitāb al-ṣaydana fī l-ṭibb* (*Pharmacology*), it has been assumed that his native language was indeed Khwarizmian.<sup>4</sup> In that work, al-Bīrūnī justifies his praise of Arabic as the scientific language par excellence by explaining that not only is Persian, which he also knows, unsuitable, but that moreover he has a mother tongue even less suitable for science, though he does not name it explicitly:

والى لسان العرب نُقلت العلوم من أقطار العالم فازدانت وحلت في الأفئدة وسرت محاسن اللغة منها في الشرايين والاوردة وان كانت كل أمة تستحلي لغتها التي الفتها واعتادتها واستعملتها في مأربها مع الألفها وأشكالها. وأقرب هذا بنفسى وهي مطبوعة على لغة لو خُلد بها علم لاستعرب استعراب البعير على الميزاب والزرافة في العراب. ثم منتقلة الى العربية والفارسية فأنا في كل واحدة دخيل ولها متكلف والهجو بالعربية أحب الي من المدح بالفارسية وسيعرف مصداق قولى من تأمل كتاب علم قد نقل الى الفارسي كيف ذهب رونقه وكسف باله واسود وجهه وزال الانتفاع به اذ لا تصلح هذه اللغة الا للأخبار الكسروية والأسمار الليلية.

“From diverse corners of the world the sciences were transferred into Arabic, were embellished, inhabited in hearts, and the niceties of the language flowed through their arteries and veins, even though each nation prefers their language, which it is familiar with and used to and uses in fulfilling its needs with its peers and familiars. *I measure this against my own self, for I was brought up in a language which, were science ever to be immortalized in it, it would be as astounding as a mule in a waterspout or a giraffe among thoroughbreds. Then I went over to Arabic and Persian, and am a stranger in each language and struggle in each one.* But I would prefer insults in Arabic to praise in Persian! He who has ever engaged with a book of science translated into Persian will know the truth of my words—how its elegance disappeared, its sense darkened, its visage blackened, and its usefulness was voided. For this language (Persian) is only fit for reciting the legends of kings or bedtime stories”.<sup>5</sup>

Reports such as al-Bīrūnī’s were already an indication that the Khwarizmian language continued to be spoken at least up until the turn of the first millennium—later in fact, than the other known Iranian languages of Central Asia, Sogdian and Bactrian, are attested in their respective homelands.

<sup>2</sup>Sachau, *Chronology*, 42. We now know that the pre-Islamic Khwarizmian writing system was derived from the Aramaic script, as were the scripts of most of the other Middle Iranian languages, see Vladimir A. Livshits, ‘The Khwarezmian Calendar and the Eras of Ancient Chorasmia’, *Acta Antiqua Academiae Scientiarum Hungaricae* 16 (1968), pp. 433–446 and more recently Pavel Jurje, ‘Some New Readings of Chorasmian Inscriptions on Silver Vessels and Their Relevance to the Chorasmian Era’, *Ancient Civilizations from Scythia to Siberia* 24, no. 1–2 (2018), pp. 279–306. Some documents in the Aramaic-derived Khwarizmian script have been discovered, but all are very difficult to read and have yet to be successfully deciphered and edited in their entirety; a number of ossuaries with Khwarizmian inscriptions are a partial exception. This article will not deal with those sources.

<sup>3</sup>Al-Bīrūnī also composed a separate treatise on the political history of Khwarizm, *Kitāb al-muṣāmara fī akhbār Khuwārizm* (*Telling Tales about the Affairs of Khwarizm*) which is lost and now known only from quotations in other works.

<sup>4</sup>For example E. S. Kennedy, ‘al-Biruni’, in *Dictionary of Scientific Biography*, vol. 2 (1981), pp. 147–158.

<sup>5</sup>Emphasis and translation mine. The introduction to the *Pharmacology* was edited and translated into German from Arabic by Max Meyerhof (‘Das Vorwort der Drogenkunde des Beruni’, *Quellen und Studien zur Geschichte der Naturwissenschaften* 3 (1932), pp. 157–208) while Hamarneh and Said (Sami K. Hamarneh & Hakim Mohammad Said, *Al-Biruni’s Book on Pharmacy and Materia Medica*, 2 vols. (Smithsonian Institution, printed under the auspices of the Hamdard National Foundation, Karachi, Pakistan, 1973–1976)) provide a handwritten edition of the entire text with translation into English and a commentary. An updated and more legible edition would be a worthwhile endeavour, especially considering the difficulty of the numerous pharmacological terms in the various languages given in Arabic transcription. The passage quoted here is based on Meyerhof’s edition, see Meyerhof, ‘Das Vorwort der Drogenkunde des Beruni’, p. 210 (German), p. 228 (Arabic), which has some irregularities. The *Pharmacology* was actually known only in its Persian version until Zeki Velidi Togan discovered the Arabic original in 1926.

### The Khwarizmian Sources

However, nothing else was really known of Khwarizmian until a series of spectacular discoveries made by the Beshkiri scholar Zeki Velidi Togan (1890–1970) between the 1920s and 1940s, which revealed two groups of Khwarizmian source material written in a modified Arabic script and recorded in Arabic texts.<sup>6</sup>

One is comprised of Islamic legal texts in Arabic containing Khwarizmian sentences. Chronologically, the first of these is a compendium entitled *Yatīmat al-dahr fī fatāwā 'ahl al-'aṣr* (*The Matchless Pearl of the Age on the Fatwas of Contemporaries*) composed by Muḥammad 'Alā' al-Dīn al-Tarjumānī al-Khuwārizmī (d. 1257), several manuscripts of which contain Khwarizmian sentences in Arabic script.<sup>7</sup> Next comes a similar type of text entitled *Qunyat al-munya li-tatmīm al-ḡunya* (*The Acquisition of that which is Desired for the Completion of the Sufficiency*), compiled in the early 13<sup>th</sup> century by Najm al-Dīn al-Zāhidī al-Ghazmīnī (d. 1260).<sup>8</sup> The *Qunya* is itself a summary of a now-lost work entitled *Munya al-fuqahā'* by the teacher of al-Ghazmīnī, Fakhr al-Dīn al-Qubaznī (known as Qāḍī Badī'), the *Qunya* itself repeating much of the material of the *Yatīma* probably via the *Munya*. Several manuscripts of the *Qunya* contain extensive text in Khwarizmian. Then, about a century later the Khwarizmian material of the *Munya* and the *Qunya* was gathered into an untitled compendium by yet another scholar of Khwarizmian origin, Jamāl al-Dīn al-'Imādī (d. ca. 1354). This latter work, otherwise untitled, has been called the *Risāla* (Treatise).<sup>9</sup>

The *Yatīma* / *Qunya* groups of texts give cases of Islamic law taken from real life in medieval Khwarizm, often including dialogue in Khwarizmian and a discussion of the extent to which utterances in Khwarizmian have the same value under Islamic law as utterances in Arabic. Composed around the 13<sup>th</sup> century CE, they show a language still in wide daily use, albeit with much borrowing from Arabic and Persian. Though undoubtedly under pressure from both, Khwarizmian appears in the texts as a still-vital language with, for example, established strategies for integrating both Arabic and Persian loans: consider the abstract noun

<sup>6</sup>As it happens, Meyerhof's work on the foreword to al-Biruni's *Pharmacology* was nearly contemporary to Togan's discovery of Khwarizmian texts, as Meyerhof notes ('Das Vorwort der Drogenkunde des Beruni', 173).

<sup>7</sup>A. Zeki Velidi Togan, 'Khwarezmische Sätze in einem arabischen Fiqh-werke', *Islamica* 3 (1927), pp. 190–213. It is essentially a critical edition of the parts of the 8 manuscripts discovered by Togan containing Khwarizmian glosses, though without translation. The sentences of this text were revisited by MacKenzie who gives a useful list of where they are repeated in the *Qunya* (David N. MacKenzie, 'Khwarizmian Enigma Variations', *Bulletin of the Asia Institute* 10 (1996), pp. 141–147). For the *Yatīmat al-dahr* see Carl Brockelmann, *History of the Arabic Written Tradition*, translated by Joep Lameer. Vol. 1, (Leiden, 2017), p. 413. This *Yatīma* is not to be confused with al-Tha'ālībī's *Yatīmat al-dahr fī mahāsīn 'ahl al-'aṣr* from more than two centuries prior.

<sup>8</sup>Togan made Walter B. Henning aware of his discovery and the latter made a brief report on the language; see A. Zeki Velidi Togan, 'Übern Sprache und Kultur der alten Chwarezmier', *Zeitschrift der deutschen morgenländischen Gesellschaft* 90 (1936), pp. \*27\*–\*30\*

and Walter B. Henning, 'Über die Sprache der Chwarezmier', *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 90 (1936), pp. \*30\*–\*34\*

<sup>9</sup>Both the *Qunya* and the *Risāla* are preserved in a single manuscript in St. Petersburg which also contains other works of al-'Imādī. The *Qunya* part of this manuscript was copied in Rajab 754 / August 1353 by a scribe named Kamāl b. 'Uḡmān b. Kamāl al-Harawī, while the *Risāla* part was copied on 4<sup>th</sup> Rajab 755 / 26 July 1354 by 'Ali b. Abī Bakr b. Ḥamid al-Ghūrī. This manuscript, which was copied from the author's autograph and corrected by him, forms the basis of MacKenzie's edition of the text, David N. MacKenzie, *The Khwarezmian Element in the Qunyat al-munya* (London, 1996), pp. 3–6. For other manuscripts of the *Qunya*, not necessarily containing Khwarizmian, see Brockelmann, *History of the Arabic Written Tradition*, p. 414.

*hl'l'wk* [halālāwak] meaning something like “halal-ness”, derived from the Arabic word *halāl* and the Khwarizmian nominal suffix *-'wk* [-āwak].<sup>10</sup>

In the 13<sup>th</sup> century, Khwarizm had long been under the rule of a succession of Turkic rulers and would be subjugated by the Mongols. Khwarizmian society was no doubt multilingual, with Arabic, Persian, and even Khwarizmian Turkic playing roles. The following extract from the *Qunya* illustrates how this text functions, how questions of language and law were considered, and the multilingual context of Khwarizm at that time. The passage first gives text in Khwarizmian (in Arabic script) and then proceeds to give an Arabic translation, as follows:

[Khwarizmian] كاس اي مرخ اي خوارزم وساخ في غشياك في زفاك في تركناك ار اُغلم اوداس هيخ كدامكام واخ  
نكور ني فاخ پا غشياكواي اخ ايخكام خي نان لفظ عتق وا كذاك كذاك

[Arabic] اي ان قال خوارزمي في اثناء التلطف على عيده بالتركية ار اُغلم و لم ينو شيء في تلطفه اُيستفاد من لفظه هذا  
العتق؟ لا.

If a Khwarizmian man says “my brave lad” [*är oğlum*] in Turkish to his slave in the course of pleasantry, and if through the pleasantry he has no intention [to manumit him] whatsoever, will manumission proceed from that word or not? No.<sup>11</sup>

The Arabo-Khwarizmian script has typically been transliterated rather than transcribed in publications by Iranists, and appears as follows in MacKenzie’s edition of the text (short vowels are not written other than when indicated by the *taškīl*, which is represented in superscript):

k’s’y mrc’y xw’rzm ws’c fy γ<sup>a</sup>šy’k fy zβ’k fy trk’nk <sup>’a</sup>r <sup>’u</sup>γ<sup>o</sup>l<sup>u</sup>m<sup>o</sup> ’wd’s hyc kd’mk’m w’c nk<sup>u</sup>wr  
ny β’c p’ γšy’k’wy<sup>a</sup>’c’ yck’m cy n’n lfz <sup>’tq</sup>w’ kδ<sup>a</sup>’k. kδ’k’.

The second group of Khwarizmian source material is found in certain manuscripts of the *Muqaddimat al-adab*, the famed Arabic dictionary of al-Zamakhsharī (d. 1144), himself also a native of Khwarizm, which have interlinear glosses in Khwarizmian. Though it was long thought that the main manuscript was his autograph, it is more probable that it dates from around 1200, nevertheless not long after the author’s death.<sup>12</sup> This material provides

<sup>10</sup>MacKenzie, *The Khwarezmian Element in the Qunyat al-munya*, p. 113.

<sup>11</sup>*Qunya*, passages 123–125 (MacKenzie, *ibid.*, pp. 17–18, 58, 188). For more about the style of these works see David N. MacKenzie, ‘Khwarezmian in the Law Books’, in *Études Irano-Aryennes offertes à Gilbert Lazard* (Paris, 1989), pp. 265–276. By “Turkish” (*bi-t-turkiyyati*) is certainly not meant modern (Anatolian) Turkish but rather one of the Turkic languages or varieties that were current in 13<sup>th</sup>-century Khwarizm.

<sup>12</sup>Togan published a facsimile of the manuscript, located in Konya, see A. Zeki Velidī Togan, *Documents on Khorezmian Culture, Pt. 1: Muqaddimat al-Adab, with the translation in Khorezmian / Horezm Kültürü Vesiklari, Kısım 1: Horezmce tercümlü Muqaddimat al-Adab* (Istanbul, 1951). An edition was later given by Johannes Benzing, *Das chwarezmische Sprachmaterial einer Handschrift der “Muqaddimat al-Adab” von Zamaxšari* (Wiesbaden, 1968), whose understanding of the Khwarizmian words was heavily critiqued by MacKenzie in a series of articles. Another copy of the *Muqaddimat al-adab* with Khwarizmian glosses was discovered in the 1980s; it was copied, according to its colophon, in 797/1395 (Nuri Yüce & Johannes Benzing, ‘Chwarezmische Wörter und Sätze aus einer choresmtürkischen Handschrift der Muqaddimat al-Adab’, *Zeitschrift der deutschen morgenländischen Gesellschaft* 135, no. 1 (1985), pp. 92–103). Henning had begun working on a Khwarizmian dictionary independently of Benzing but only the initial part was ready and was only published posthumously, Walter B. Henning, *A Fragment of a Khwarizmian Dictionary*, edited by D. N. MacKenzie (London, 1971).

the majority of the known Khwarizmian lexicon with over 4,000 glosses, often individual words or brief phrases rather than sentences.

Taken together, all these sources shed light on Khwarizmian as it was written in the 12<sup>th</sup> and 13<sup>th</sup> centuries in the Arabic script. And indeed, the manuscripts' relative consistency in spelling and use of new letter-forms leads one to assume that they in fact are written in a roughly 'standard' Arabo-Khwarizmian script, even if later copyists did not always adhere to it or understand the Khwarizmian. And if, as al-Bīrūnī mentions, the Arab conquest ultimately led to the loss of knowledge of the indigenous script, it would be unsurprising if an Arabo-Khwarizmian one had developed soon after. Just as the Arabic script was extended to represent certain sounds required for Persian, such as ڄ for [č] and پ for [p], at some point it was also extended to represent Khwarizmian, in particular by the innovation of a new letter: ځ, a *hā'* with three dots on top.<sup>13</sup> This ځ of the texts has been transliterated with a *c* by Iranists, as can be seen in the above extract from the *Qunya*. This goes back to Henning, who proposed transliterating *c* for ځ on the basis of both modern Pashto where the letter ځ represents [ts] and modern Ossetic in which Iranian \*č and \*-ti- become [ts], a sound change which Henning proposed also for Khwarizmian.<sup>14</sup> Later, Henning proposed that *c* encodes both this [ts] and a voiced allophone [dz]—also as in Pashto, where ځ was used for [ts] and [dz] from the late 17<sup>th</sup> century until the 20<sup>th</sup> century when the separate sign ځ (a *hā'* with *hamza* above) was developed for [dz].<sup>15</sup>

Khwarizmian has a [č] ڄ besides this ځ, a distinction made quite consistently in the *Qunya/Risāla*, though somewhat irregularly in the *Muqaddima*.<sup>16</sup> The conditions under which both sounds occur have not been sufficiently clarified, however.

Since the *Qunya* is rather consistently pointed, it is possible to see that [č] ڄ occurs primarily in Persian loanwords (such as č'h 'pit' from Persian čāh) but also in inherited Khwarizmian words as a secondary change from earlier consonant clusters (such as 'čn 'to be thirsty' < \*tršn-). There are a handful of words written with č where the č may go back to Old Iranian \*č (such as čm 'skin' < \*čarman-, but this could also be a Persian loan), but the majority of words with older \*č are written with the ځ (such as cm 'eye' < \*čašman- or q'r

<sup>13</sup>Khwarizmian also uses a *fā'* with three dots above to represent [v], seemingly a distinct phoneme from [w]. This letter occurs also in very early New Persian manuscripts as a way of indicating [β] since the *wāw* already indicated [v]. It is thus not uniquely Khwarizmian, but the fact that they are shared is suggestive. For more on this letter and on the development of Arabo-Persian orthography in general, see now Paola Orsatti, 'Persian Language in Arabic Script: The Formation of the Orthographic Standard and the Different Graphic Traditions of Iran in the First Centuries of the Islamic Era', in *Creating Standards: Interactions with Arabic Script in 12 Manuscript Cultures* (Studies in Manuscript Cultures 16), (ed.) D. Bondarev, A. Gori, and L. Souag (Berlin, 2019), pp. 39–72.

<sup>14</sup>Henning 'Über die Sprache der Chvarezmier', p. \*31\*.

<sup>15</sup>Walter B. Henning, 'Mitteliranisch', in *Handbuch der Orientalistik* 1.4.1 (Leiden, 1958), pp. 20–130 [82]. The first attested use of ځ for [ts] and [dz] in Pashto is from 1697, whereas the earliest manuscripts (1651–1690) employ ځ for [ts] and a *dāl* with subscript dot for [dz], cf. David N. MacKenzie, 'A standard Pashto', *Bulletin of the School of Oriental and African Studies* 22, no. 1–3 (1959), pp. 231–235 [233]. In Pashto the affricates [ts] and [dz] derive from the depalatalisation of older \*č and \*j, respectively. In Khwarizmian, older \*j as well as \*ž seem to become [z].

<sup>16</sup>The manuscripts of the *Muqaddima*, particularly the one with the most extensive Khwarizmian glosses edited by Benzing (*Das chwaresmische Sprachmaterial*), use not only ځ "underpointed" to ځ but also ڄ underpointed to ڄ for words which seem on the basis of their occurrence elsewhere or their etymology to have ځ. This makes the task of establishing the possible difference between [ts] and [č] in native Khwarizmian difficult. Additionally, a few words in the *Muqaddima* are pointed as ځ and ڄ simultaneously for an unknown reason: *'awn* is pointed as ځ and ڄ (Henning, *A Fragment of a Khwarizmian Dictionary*, p. 13), and *'fč'wy* is ځ and ڄ (*ibid.*, p. 16). Although the first of these is suggestive of an attempt at indicating a voiced sound, perhaps it is merely decorative as in the same manuscript simultaneous pointing is also found on *'δr* (one point under and one over the *dāl*), *nm 'sry* (*n* and *b*), and *zwz* (*z* and *ž*).

Table 1. Extended Arabic letter-forms for Khwarizmian in the *Risāla*

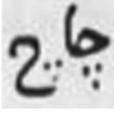
	
<i>mrc</i> 'man' (fol. 241v20)	č'č 'yam' (fol. 248r7)

Table 2. Examples of palatalisation to *c* in Khwarizmian drawn from the *Qunya* and *Muqaddima*

* <i>t</i> > <i>c</i>	* <i>k</i> > <i>c</i>	* <i>d</i> > <i>c</i> (mostly after * <i>n</i> ?)
* <i>martiya-</i> > <i>mrc</i> 'man'	pl. *- <i>ki</i> > <i>-c(i)</i>	* <i>bandaya-</i> > <i>βncy-</i> 'to tie'
* <i>ati-iyā-</i> > <i>cy-</i> 'to enter'	* <i>zānūki</i> > <i>z'nuč</i> 'knees'	* <i>pati-bandaya-</i> > <i>pcβncy-</i> 'to connect'
	f.gen.sg. *- <i>kiya</i> > <i>-c(a)</i>	
	* <i>γōkiya</i> > <i>γwč</i> 'of the cow'	

'four' < čaθwāra-); the conditions under which \*č is preserved as [č] have yet to be fully established. Several other sound changes in Khwarizmian have evidently led to *c* being a frequently occurring letter: these include the palatalisation of \**t*,<sup>17</sup> the palatalisation of \**k*, and also seemingly the palatalisation of \**d*, in the environment of high vowels or the palatal glide [y].

As mentioned, it was words such as *βncy* which led Henning to suggest that č not only encoded a voiceless affricate [ts] but also a voiced counterpart [dz] deriving from earlier \**d*; many if not most of these cases involve the sequence \*-*nd-*, one of the few places where \**d* did not change into a fricative [ð] as it does elsewhere. We shall return to this discussion. In any case, this č seems to have been invented specifically for the needs of Khwarizmian and these manuscripts represent the earliest attestation of the letter č, at least three centuries before it is used for Pashto for the first time. Yet when it was first used for Khwarizmian cannot be said with any certainty. Manuscripts of al-Bīrūnī's works in which he cites Khwarizmian terms do not employ the č, perhaps because they were copied by later, non-Khwarizmian-speaking scribes who did not know of the letter—in the Edinburgh manuscript (copied 1307) of the *Chronology*, the ultimate source of Sachau's manuscripts, words with *c* are written with either چ or چ, and in the Beyazit manuscript, the oldest (12<sup>th</sup> c.) and best manuscript of the *Chronology*, such words are written with either no points

<sup>17</sup>The palatalisation of *t* [t] to *c* [ts] does not seem to have taken place yet in the attested pre-Islamic Khwarizmian texts, or, if it had, the older writing system based on the Aramaic script did not represent the change after it had taken place in the spoken language. Moreover, it is worth pointing out, as MacKenzie already noted, that "there are many examples of differing developments of certain sounds, suggesting either a mixture of dialects or the adoption of loanwords from several neighbouring languages" (David N. MacKenzie, 'Chorasnia III: The Chorasnia Language' in *Encyclopedia Iranica* (online edition, 1991/2011), url: <http://www.iranicaonline.org/articles/chorasnia-iii> (last accessed 24 December 2019). It is a possible scenario that the redactors of the 13<sup>th</sup>-century texts gathered examples from informants from different areas with knowledge of different varieties.

(ح) or just one (ح)<sup>18</sup>—or because it had not been invented yet for the language. The earliest Khwarizmian source in Arabic script, the Konya manuscript of the *Muqaddimat al-adab*, does not appear until perhaps the end of the 12<sup>th</sup> century. But as with most ancient languages which are known today only in written form, it is difficult to know *exactly* how certain sounds were pronounced, and this ح is no exception. Fortunately, though, there exists a contemporary source potentially able to shed some light on the matter.

### Ibn Sīnā's Remarks on Khwarizmian

Written between his arrival at the Iṣfahān court of 'Alā' al-Dawla in 1024 and his death in 1037, Ibn Sīnā's treatise *Asbāb Ḥudūt al-Ḥurūf* (*The Causes of the Genesis of the Consonants*) gives a rigorous treatment of Arabic phonetics, detailing the various sounds in the Arabic language and the parts of the mouth involved in producing them.<sup>19</sup> While the treatise draws on ancient traditions, especially Galen, it also contains unique and novel arguments about the physical production of sound, no doubt based on Ibn Sīnā's medical expertise. Several features of the work, from the order in which the letters are discussed to the linguistic terminology to several of the concepts (such as *qal'* "sudden separation" and *ruṭūba* "moisture"), set it apart from the classical Arabic linguistic traditions.<sup>20</sup>

In addition, the treatise has a chapter entitled *fi l-ḥurūf al-ṣabīha bi-hādihihī l-ḥurūf wa-laysat fi lughat al-'arab* "Regarding consonants similar to these [Arabic] consonants but not in the language of the Arabs",<sup>21</sup> in which are discussed both Arabic consonants which are produced incorrectly by non-Arabs, as well as sounds that were not part of Arabic but occurred in other languages with which he was familiar. His method involves comparing these sounds to the Arabic consonants he describes earlier in the treatise. For example, Ibn Sīnā notes that other languages have "ḡīm-like" consonants, "among them [being] the consonant which is pronounced at the beginning of the noun 'well' in Persian, which is *čāh*" (*minhā*

<sup>18</sup>My thanks to François de Blois for the information about the pointing in these manuscripts; de Blois adds (personal communication) that pointing and vocalisation in the Edinburgh manuscript is largely decorative. Henning ('Mitteliranisch', 83) suggests that al-Bīrūnī employed an already-existing Arabic orthography for Khwarizmian in his *Chronology*, but I do not see how this can be proven.

<sup>19</sup>See Dimitri Gutas, *Avicenna and the Aristotelian Tradition: Introduction to Reading Avicenna's Philosophical Works*, 2nd edition (Leiden, 2014), pp. 442–443, and Jules L. Janssens, *An Annotated Bibliography on Ibn Sīnā (1970–1989)* (Leuven, 1991), p. 49, for information about the extant manuscripts of the *Asbāb Ḥudūt al-ḥurūf*, sometimes known incorrectly under the title *Makhārij al-ḥurūf*. Gutas points out that as it was composed at the request of a philologist named Abū Manṣūr Muḥammad al-Jabbān at the court of 'Alā' al-Dawla in Iṣfahān, it must have been composed after 415/1024. On Ibn Sīnā's time in Iṣfahān, see Lenn Goodman, *Avicenna* (London, 1992) pp. 30ff. The *Causes* has been edited several times and translated into English twice: Khalil Semaan, *Arabic Phonetics. Ibn Sīnā's Risālah on the Points of Articulation of the Speech-sounds translated from Medieval Arabic* (Lahore, 1963) only gives an English translation without commentary or Arabic text while Solomon Sara, *A Treatise on Arabic Phonetics. Translation, Notes & Comments* (Munich, 2009) is a fuller study, including Arabic text and commentary. The Arabic text used in this article is that of Sara's edition—note that all translations from this text are mine with reference to Sara's translation.

<sup>20</sup>Important analyses of the theories which Ibn Sīnā develops behind these two terms are István Ormos, 'Observations on Avicenna's Treatise on Phonetics', *Acta Orientalia Academiae Scientiarum Hungaricae* 39, no. 1 (1985), pp. 45–84 and *id.*, 'A Key Factor in Avicenna's Theory of Phonation', *Acta Orientalia Academiae Scientiarum Hungaricae* 40, no. 2–3 (1986), pp. 283–292. For other studies on this text not cited here, see Janssens, *An Annotated Bibliography on Ibn Sīnā*, pp. 153–155.

<sup>21</sup>One version of the treatise has the shorter chapter title *fi l-ḥurūf al-ṣabīhati bi-hādihihī l-ḥurūf*.

*l-ḥarfū lladhī yuntaqu bihi fī awwali smi l-bi'r bi-l-fārisiyyati wa-huwa ḥāh*).<sup>22</sup> Because of properties such as its place of articulation, Persian [č], Ibn Sīnā notes, and correctly from a modern linguistic perspective, that it is similar to the Arabic [ġ]. He then states that there are other sounds which are not in Arabic or Persian but in other languages, such as a “ṣād-like” (*šibh al-ṣād*) consonant and a “sīm-like” (*šibh al-sīm*) consonant, for which he neither specifies the language in which they occur nor gives any examples. He then goes on to describe a “zāy-like sīm that is frequent in the language of the people of Khwarizm” (*sīmun zā'iyyatun takṭaru fī lughati 'ahli khuwārizm*).<sup>23</sup>

Before coming to a discussion of this sound, it is worth asking whether Ibn Sīnā was referring to that which we now know of as the Khwarizmi language, as opposed to a distinctive local variety of (New) Persian, since he says the “language of the people of Khwarizm” (*lughat 'ahl khuwārizm*) while the somewhat later sources discussed in the first part of this paper use “Khwarizmi” (*khuwārizmī*); these latter sources were, of course, written by actual speakers who no doubt knew what to call their own language, even in Arabic. Scholars of the generation just prior to Ibn Sīnā were aware of, or had encountered, a distinct language in the region, though for the most part they did not give it a specific name: the geographer al-Maqdisī (d. 991) simply mentions that the “language of the people of Khwarizm cannot be understood” (*lisān 'ahl khuwārizm lā yufhim*)<sup>24</sup> while the noted traveller Ibn Faḍlān (d. 960) was somewhat more judgmental, writing in his travelogue that “the Khwarizmi are the most barbarous of people, both in speech and in custom. Their speech sounds like the cries of starlings (*kalāmuhum 'aṣbaha ṣay'in bi-ṣiyāḥi z-zarāzīr*). There is a village... whose inhabitants are known as Kardaliya, and their speech sounds like the croaking of frogs (*kalāmuhum 'aṣbahu ṣay'in bi-naḥīqi ḍ-ḍafādi*)”.<sup>25</sup> Ibn Ḥawqal (d. ca. 978), who was in Khwarizm in 969, was more objective, stating that “[the Khwarizmi] language is unique to them, no other like it is spoken in Khurāsān (*wa-lisān 'ahlīhā mufrad bi-luḡatihim wa-laysa bi-khurāsān lisān 'alā luḡatihim*)”.<sup>26</sup> So well before even al-Bīrūnī wrote about it, scholars of the time seem to have been aware of a particular and seemingly unique language in the region, and this general knowledge is likely to have been available to Ibn Sīnā.

More importantly, however, Ibn Sīnā spent about a decade, until 1012, living and working in Khwarizm at the court of the Khwarizm-Shāhs at Gurgānj, where he undoubtedly heard Khwarizmi being spoken and actually overlapped with al-Bīrūnī, with whom he also corresponded later in life.<sup>27</sup> In fact, like his polymath colleague, Ibn Sīnā may also even have been a speaker of a non-Persian Iranian language before learning and mastering both Persian and Arabic.<sup>28</sup> Given all this, it seems certain that Ibn Sīnā is indeed referring to the Khwarizmi language known to us.

<sup>22</sup>Sara, *A Treatise on Arabic Phonetics*, p. 37. Sara's commentary to this chapter of the work is quite brief and does not attempt to compare Ibn Sīnā's descriptions with data from the languages he alludes to (Persian and Khwarizmi).

<sup>23</sup>Sara, *A Treatise on Arabic Phonetics*, pp. 37–39, 72–73. Sara seems to be unfamiliar with the Khwarizmi language, calling it the “dialect of Khwarizm, area north east of Tehran” (*ibid.*, p. 121).

<sup>24</sup>Basil Anthony Collins, translator, *The Best Divisions for Knowledge of the Regions* (Reading, 1994), p. 272

<sup>25</sup>Ibn Faḍlān, *Risālat Ibn Faḍlān*, edited by Sāmī Dahhān (Damascus, 1959), p. 82.

<sup>26</sup>Ibn Ḥawqal, *Kitāb ṣūrat al-arḍ*, edited by J. H. Kramer (Leiden, 1938), vol. 2, pp. 477–478, 481–482.

<sup>27</sup>For details of Ibn Sīnā's time in Khwarizm, see Goodman, *Avicenna*, pp. 19–24. Goodman refers to al-Bīrūnī, strangely, as a “Khwarizmi Persian”.

<sup>28</sup>For Ibn Sīnā's background, see Goodman, *Avicenna*, 11. Sara (*A Treatise on Arabic Phonetics*, p. 36 n. 75) states that Ibn Sīnā was a native speaker of Persian, but it would be surprising if in the late 10<sup>th</sup> century the entire region

Now, the entire passage to be discussed is as follows:

ومن ذلك سين زائبة تكثر في لغة اهل خوارزم وتحدث بأن تهيأ الهيئة التي عن مثلها تحدث السين ثم يحدث في العضلة الباطحة للسان ارتعاد كما يحدث في الزاي يلزم ذلك الارتعاد مماستات خفية غير محسوسة يحتبس لها الهواء احتباساتٍ غير محسوسة فتضرب السين لذلك الى مشابهة الزاي.

Among these [sounds not occurring in Arabic] is a *sīn zā'iyya* that is frequent in the language of the people of Khwarizm. It occurs by preparing the construction from the like of which the *sīn* occurs, then in the flattening muscle of the tongue an *irti'ād* occurs, as occurs with *zāy*. That *irti'ād* is accompanied by hidden, imperceptible contacts, by which the air is trapped by imperceptible obstructions (*ihtisābāt*). Thereupon the *sīn* occurs like the *zāy*.<sup>29</sup>

How the *sīn* and *zāy* are to be combined is at first glance difficult to envisage. Not so for the next sound described in the chapter, however, which is a *šīm*-like *zāy* (*zāy šīmiyya*) of the kind, Ibn Sīnā says, heard in Persian when they say *šarf* 'deep' (*zāyūn šīmiyyatun tusma'u fi l-lughati l-fārisiyyati 'inda qawlihim šarf*).<sup>30</sup> The point is quite clear: the *zāy* pronounced at the place of articulation of the *šīm* gives us [ž].<sup>31</sup> In modern linguistic terms, adding the voicing of the fricative [z] to the palato-alveolar articulation of the fricative [š] yields the voiced palato-alveolar fricative [ž]. Its writing with *ž* in the Arabo-Persian script, it is worth noting, likewise points to its association with Arabic *ج* rather than with *ج*. But if a [š]-like [z] is the sound [ž], then [z]-like [s] is not a new sound but simply remains [z]. That is, adding of the voicing of [z] to [s] just gives [z]. We might try to match the sound described in the *Causes* with what we already think we know of Khwarizmian from the sources previously mentioned. There are several *zāy*-like and *sīn*-like sounds which are potential candidates for what Ibn Sīnā describes: besides [s] and [z] themselves, there are also [š] and [ž], as well as [č] and the sound written by means of *خ*. We can firstly eliminate [s], [z], and [š] from the list, as they, also occurring in Arabic, would not have merited any special comment by Ibn Sīnā. We can also eliminate [č] since he describes that separately from the *sīn zā'iyya*, as mentioned. What then could the sound be? Since Ibn Sīnā unfortunately cites no example from the Khwarizmian language, we must interpret his description of the consonants to determine what sound he understands this *sīn zā'iyya* to be. First is that the *sīn zā'iyya* is based on the construction of the *sīn*, the description of which is rather concise:

وأما السين فتحدث مثل حدوث الصاد إلا أن الجزء الحابس من اللسان فيه أقل طولاً و عرضاً وكأنها تحبس العضلات التي في طرف اللسان لا بكليتها بل بأطرافها

from Balkh to Bukhārā spoke only New Persian. In fact, al-Bīrūnī frequently cites words from “Bukhārān” (*al-bukhāriyya*) in his *Pharmacology* next to Sogdian and several other languages; the examples given suggest that “Bukhārān” was very close to, if not a variety of, Sogdian, cf. Henning, ‘Mitteliranisch’, p. 85. Al-Maqdisī also noted that the language of Sogdiana was similar to that of the rural districts of Bukhārā (*wa-li-š-šuyūd lisān 'alā hida yuqāribuha alsinat rasātiq bukhārā*), but his examples of Bukhārān speech seem to be simply Persian, cf. Collins, *The Best Divisions for Knowledge of the Regions*, p. 273.

<sup>29</sup>Sara, *A Treatise on Arabic Phonetics*, pp. 36–39, 72–73.

<sup>30</sup>*Ibid.*, pp. 38–39.

<sup>31</sup>This method of describing a sound not occurring in Arabic, or at least not represented in the Arabic script, by citing two Arabic sounds which combine to yield something close to it may derive in part from an earlier work. Ḥamza al-Iṣfahānī (d. after 961), in his *al-Tanbīh 'alā ḥudūd al-taṣṣīf*, mentions several Persian consonants, such as the [p] which is “between the *fā'* and the *bā'*” (*bayna l-fā' wa-l-bā'*), see al-Iṣfahānī, *al-Tanbīh 'alā ḥudūd al-taṣṣīf*, 2nd edition, edited by Muḥammad As'ad Ṭālas (Beirut, 1992), pp. 34–35. Notably, his description of a few consonants corresponds to what is known of late Middle Persian or Early New Persian phonology.

As for the *sīn*, it occurs like the occurrence of the *ṣād* except that the obstruction (*ḥabs*) of the tongue in it is less in length and in width. It is as though the muscles that are at the edge of the tongue obstruct (*taḥbis*) not with their entirety but with their edges.<sup>32</sup>

The *sīn* is thus related to the *ṣād* in terms of its “obstruction” (*ḥabs*), but is said to be less (*aqall*) and to not involve the entirety of the tongue. If we turn to the *ṣād*, we find that it is likewise described in relation to the *sīn*, where it is said to have a “narrower” (*adyaq*) and “drier” (*aybas*) obstruction than the *sīn* but that it covers (*yutbiq*) two-thirds of the surface of the palate. If Ibn Sīnā’s description of the *sīn* is brief, though, then his description of the *zāy* is much more long and complicated, but worth citing in full:

وأما الزاي فإنها تحدث من الأسباب المصفرة التي ذكرناها إلا أنّ الجزء الحابس فيها من اللسان يكون ممّا يلي وسطه ويكون طرف اللسان غير ساكن سكونه الذي كان في السين بل يمكن من الاهتزاز فإذا انفلت الهواء الصافر عن المحبس اهتز له طرف اللسان واهتزت رطوبات تكون عليه وعنده ونقص من الصفير . إلا أنه باهتزازة يحدث في الهواء الصافر المنفلت شبه التدرج في منافذه الضيقة بين خلل الأسنان فيكاد أن يكون فيه شبه التكرير الذي يعرض للراء وسبب ذلك التكرير اهتزاز جزء من سطح طرف اللسان خفي الاهتزاز.

As for the *zāy*, it occurs from the whistling causes that we mentioned except that the *ḥabs* of the tongue emerges near its middle and the edge of the tongue is not holding the stationary position that occurs in the [articulation of the] *sīn*, but is, rather, capable of *ihtizāz*. If the whistling air escapes the place of obstruction (*maḥbas*), the edge of the tongue vibrates (*ahtazza*) to it; the moistures that it has and that are on it vibrate (*ahtazza*) and it has a diminishment of the whistling, except that in its *ihtizāz* it causes in the whistling and coursing air a quasi-tumble in its narrow passages between the gaps of the teeth. There is in it almost the quasi-repetition that happens to the *rā*’ and the cause for that repetition is the *ihtizāz* of a part of the surface of the edge of the tongue with a hidden *ihtizāz*.<sup>33</sup>

The first way in which the *sīn* *zā’iyya* is then *zāy*-like is that, according to its description, “in the flattening muscle of the tongue an *irti’ād* occurs, as occurs with *zāy*” (*yaḥduṭu fī l-‘aḍalati l-bāṭiḥati li-l-lisāni kamā yaḥduṭu fī l-zāy*). The description of the *zāy*, however, mentions no *irti’ād*, which I have left untranslated for the moment. Instead, the *zāy* is described as differing from *sīn* with regards to *ihtizāz*: “the edge of the tongue is not holding the stationary position that occurs in the [articulation of the] *sīn* but is, rather, capable of *ihtizāz*” (*yakūnu ṭarafu l-lisāni ḡayri sākinin sukūnahū lladhī kāna fī l-sīni bal yumkinu mina l-ihtizāz*). The related consonant *zāy sīniyya* [ž] is also described as “manifested by the *ihtizāz* of the surface of the tip of the edge of the tongue” (*tu’riḍu bi-htizāzi saḥhi ṭarafī l-lisān*). So, to the *zāy* itself and the *sīn*-like *zāy*, Ibn Sīnā ascribes the characteristic of *ihtizāz*, which he only uses for the small group of consonants presented in Table 3 below.

As can be seen, this *ihtizāz* or *irti’ād* is employed only for the Arabic consonants *zāy* [z] and *ḍāl* [ḍ], and for several non-Arabic consonants such as the “*fā*’ which almost resembles the *bā*” (by which is meant the Persian [v] or [β]), the *sīn*-like *zāy* (Persian [ž]), the *zāy*-like *sīn* of Khwarizmian, and a *zā*’-like *zāy* in an unspecified non-Arabic language. Ibn Sīnā only otherwise uses the term in the description of the *ḡayn*, where he says the airflow causes something similar to *ihtizāz* (but not *ihtizāz* itself). Since the description mostly uses *ihtizāz*,

<sup>32</sup>Sara, *A Treatise on Arabic Phonetics*, pp. 30–31.

<sup>33</sup>Sara, *A Treatise on Arabic Phonetics*, pp. 30–31.

Table 3. Consonants to which Ibn Sīnā ascribes *iẖtizāz* or *irti'ād*

Consonant	Description	Translation
<i>zāy</i>	<i>yakūnu ʔaraʔi l-lisāni ġayri sākinin sakūnahū llađi kāna fi s-sīn bal yumkinu mina l-iẖtizāz</i>	“the edge of the tongue is not holding the stationary position that occurs in the [articulation of the] <i>sīn</i> but is, rather, capable of <i>iẖtizāz</i> ”
<i>đāl</i>	<i>bi-mā yađi ʔarfa l-lisāni mina l-ruḡbati ḥattā yuḥarrakahā wa-yahuzzuḥā hazzan yašīran ... yakūnu fi đ-đāl qaḡban mina l-iẖtizāzi llađi yakūnu fi z-zāy</i>	“with what follows the edge of the tongue of moisture so that it moves it and oscillates it ( <i>yahuzzuḥā</i> ) slightly ... the <i>đāl</i> , it is closer with respect to the <i>iẖtizāz</i> that is present in the <i>zāy</i> ”
<i>sīn zā'yyya</i>	<i>yaḥduḡu fi l-'ađalati l-bāḡiḥati li-l-lisān irti'ādun ka-mā yaḥduḡu fi z-zāy</i>	“an <i>irti'ad</i> occurs in the flattening muscle of the tongue like that which occurs in the <i>zāy</i> ”
<i>zāy sīniyya</i>	<i>wa-ḥiya sīnun ... ta'riđu bi-htizāzi saḡḡi ʔaraʔi l-lisān</i>	“it is a <i>šīn</i> ... manifest by an <i>iẖtizāz</i> of the surface of the edge of the tongue”
<i>zāy zā'yyya</i>	<i>yakūnu wasaḡu l-lisāni fiḥā 'arfa'a wa-l-iẖtizāzu fi ʔaraʔi l-lisāni xaḡiyyun jiddan</i>	“in [the <i>zā</i> '-like <i>zāy</i> ] the center of the tongue is higher and the <i>iẖtizāz</i> at the edge of the tongue is very hidden”
<i>fā' tukād</i> <i>tušbahu l-bā'</i>	<i>ḥā-hunā fā' tukādu tušbihu l-bā' ... yukādu yaḥduḡu minhu fi s-saḡḡi llađi fi bāḡini š-šifati htizāz</i>	“here is a <i>fā'</i> that almost resembles the <i>bā'</i> ... an <i>iẖtizāz</i> almost occurs on the inner surface of the lip”

and only once *irti'ād* (which is linked to the *zāy* where *iẖtizāz* is used), it is possible that both terms are meant to refer to the same phenomenon. In particular, this seems to be a characteristic of what we would now call *voiced fricatives*. Though in 2009 Sara translated *iẖtizāz* as “oscillation” but *irti'ād* as “trembling”, implying a difference in the two, I think that on the basis of their usage and the similarities in the consonants grouped above, it is reasonable to assume that Ibn Sīnā intended them to describe the same phenomenon.<sup>34</sup> As a property common to voiced fricatives, both terms may best be translated with “vibration”. Yet in general the *Causes* does not make use of a category comparable to our modern category of “voice”. Instead, by using “vibration”, the treatise tends to point to where vibration, as an effect of voiced consonants, can be felt in the mouth: for [z] and [ð] it is on the “edge of the tongue” (*ʔaraʔi l-lisān*) while for the *sīn zā'yyya* it is in the “flattening muscle of the tongue” (*al-'ađalati al-bāḡiḥati li-l-lisān*), for the [ž] it is on the “surface of the edge of the tongue” (*saḡḡi ʔaraʔi l-lisān*), and for the [v] or [β] it almost occurs on the “inner (surface) of the lip” (*bāḡini š-šifati*).

The second characteristic of the *sīn zā'yyya* is that the airflow is trapped by “imperceptible obstructions” (*iḥ̱tisābāt ġayr maḡsūsa*). The feature of “obstruction” (*ḥabs*) occurs frequently in the work and appears to be a fundamental feature of Ibn Sīnā's phonetic description. *Ḥabs* is used to describe how and where oral elements touch each other to change the airflow and produce different sounds: this could be the tongue touching the palate (as in the *šād*), but could also be both lips touching each other (as in the *bā'*). Many consonants have a “complete obstruction” (*ḥabs tāmm*), some have an “incomplete obstruction” (*ḥabs ġayr tāmm*), and

<sup>34</sup>Ormos (“Observations on Avicenna's Treatise on Phonetics”) points out that the terminology used by Ibn Sīnā in this work is not entirely systematic and consistent, though he does not mention *iẖtizāz* specifically.

others, interestingly enough, are described as having no *ḥabs*, in particular the *ṣīn*, which is said to be like the *jīm* but with no obstruction at all.<sup>35</sup> Since the *ḥabs* is what is involved in obstructing the oral cavity, it does not just have quantity (“complete” or “incomplete”) but also quality: as we have seen it may be “less” (*aqall*) as in the *sīn* or “drier” (*aybas*) as in the *ṣād*.<sup>36</sup>

Do these descriptions help us to understand the articulation of the *sīn zā’iyya*? First, it is related to the *sīn* (and thereby also to the *ṣād*) in the quality and quantity of its *ḥabs*. This suggests that among obstruents, it belongs towards the fricative (including [s] but excluding [ʃ]) to affricate side of the consonant group (the only affricate in Ibn Sīnā’s system seems to be [ʃ]). Secondly, it has “vibration” like the *zāy* and other voiced fricatives. Ibn Sīnā’s system thus suggests that the *sīn zā’iyya* is a voiced alveolar sibilant fricative or affricate.

As mentioned previously, Iranistic scholarship has postulated the sounds [ts] and [dz] for Khwarizmian, both encoded by the Arabic letter *ṣ*. If our analysis of Ibn Sīnā’s description of the *sīn zā’iyya* is correct, then the best match for it seems to be [dz] rather than [ts]. From the perspective of our understanding of the Khwarizmian sources, this is somewhat unexpected, as [ts] seems to be the more common sound, at least on the basis of etymology. One possibility is that what has been thought thus far to be a [ts] was actually a voiced [dz], and this [dz] was therefore one of the most prominent “foreign” sounds to an Arabic or Persian ear. It seems odd that Ibn Sīnā, with his firsthand knowledge of Khwarizmian and ability to describe both differing pronunciations of Arabic sounds such as *qāf*<sup>37</sup> and non-Arabic sounds such as the [p], [v], [č], and [ž] of Persian, would not have been able to notice both a [ts] and a [dz], if both existed. But his treatise on phonetics does not cover all the possible sounds he could have heard in the various languages spoken in the places he lived; perhaps a [dz] was more unusual to him than other sounds and therefore merited description.<sup>38</sup>

At the same time, Ibn Sīnā does not claim to exhaustively describe *all* non-Arabic sounds he had ever heard, and does not consistently give examples for those which he does describe. It thus seems that the section on non-Arabic sounds in the *Causes* most likely served to give further examples illustrating the applicability of his phonetic approach to speech sounds in general. Ibn Sīnā’s remarks do not, unfortunately, bring a definitive solution to our study of this aspect of Khwarizmian phonology. It is nevertheless my contention that he was

<sup>35</sup>Sara, *A Treatise on Arabic Phonetics*, p. 29.

<sup>36</sup>See Sara, *ibid.*, pp. 90–92 for more discussion of *ḥabs* in this work. In fact, it corresponds quite closely to what we would call “obstruents”, a category based on whether speech sounds obstruct airflow.

<sup>37</sup>See Haïm Blanc, ‘Les deux prononciations du qāf d’après Avicenne’, *Arabica* 13, no. 2 (1966), pp. 129–136, who attempts to connect the two pronunciations of *qāf* in the *Causes* to modern Arabic dialects.

<sup>38</sup>It is an intriguing but perhaps unlikely possibility that *ṣād* was still pronounced as an affricate in Ibn Sīnā’s time and that he therefore would not have thought a voiceless affricate unusual in Khwarizmian vis-à-vis Arabic but would have noticed a voiced one; Ahmad Al Jallad (‘Aṣ-ṣādu llaṭī ka-s-sīn: evidence for an affricated ṣād in Sibawayh?’, *Folia Orientalia* 51 (2014), pp. 51–57) has recently argued that *ṣād* was still an affricate [tʃ] at the time of Sibawayh, but this is two centuries before Ibn Sīnā. From the description of the *ṣād* in the *Causes*, it is difficult to tell if Ibn Sīnā is describing an affricate. The key phrase may be “the air escapes from that *maḍīq* after a great deal of it has been blocked from behind” (*wa-yatasaraba l-hawā’u ‘an dālīka l-maḍīqi ba’da ḥaṣri ṣay’in kaḥrīn min-lu min warā’*) (see Sara, *A Treatise on Arabic Phonetics*, pp. 28–31), but the phrase is ambiguous in that it employs *ḥaṣr*, a term otherwise not occurring in the text, perhaps for *ḥabs*, and in that it could simply be describing the fact that the raised tongue restricts the oral cavity a great deal, especially in comparison with the *sīn*, to which it is related in the *Causes*.

describing a real sound present in the Khwarizmian language of his day, and quite possibly the sound, or one of the sounds, encoded by the Arabic letter  $\text{ع}$ , developed specially for writing Khwarizmian. What scholars reconstruct of the phonology of a medieval language preserved only in written texts is, of course, tentative. It is thus instructive to consider examples, rare as they may be, of linguistic analysis of such languages dating from a time in which they were still spoken. That the two may not neatly agree is an invitation both to revisit our understanding of those texts and to continue to cast our nets wider in search of contemporary sources.

ADAM BENKATO  
*University of California, Berkeley*  
[abenkato@berkeley.edu](mailto:abenkato@berkeley.edu)