Introduction

إِنَّا أَنْزَلْنَاهُ قُرْآنًا عَرَبِيًّا لَعَلَّكُمْ تَعْقِلُونَ

The Quran, Q12:2, in the reading of Ibn Kaṭīr

∴

1.1 Previous Scholarship

The main question the current book aims to answer is: “What is the language of the Quran?” Despite more than a century of in-depth study of the Quran, and a debate on the linguistic nature, I believe the discussion on this question has not progressed significantly, since Vollers (1906). Despite the many deficiencies of Vollers’ work, which have already been addressed in detail especially by Nöldke (1910), and Geyer (1909), I believe that the question it formulated remains essential to furthering our understanding of the linguistic history of Arabic and the context in which the Quran arose: What is the language of the Quran and, perhaps more importantly, *how do we know?*

Vollers’ radical theory proposed that the Quran was originally composed in the Hijazi common language (*Volksprache*)—a language he considered much more akin to modern Arabic dialects than the literary language (*Schriftsprache*) in which the Quran later came to be recited. He believed it was reworked by Arab grammarians to adhere to the Classical standard of this literary language, the *ʿarabiyyah*. He saw the literature on the Quranic reading traditions (*Qirāʾāt*) and the variants described within them as providing essential traces of the “original” language. From this he concluded that the Quran was reworked in later times by Arab grammarians, making massive changes to the linguistic nature of the text, including pervasive changes to the consonantal skeleton of the text. Especially the supposition of relatively late changes to the consonantal skeleton by grammarians has become difficult to accept. It is now certain that the standardization of the text well preceded the work of the Arab grammarians by over a century (Sinai 2014a; 2014b; van Putten 2019c).

But, while Vollers’ theory has several shortcomings that make it not quite convincing today, the fundamental question as to what the language of the

Quran was, is still a valid one, and it has never adequately been answered. Instead, a consensus developed although no evidence seems to have ever been proffered for it. As Rabin (1955, 24) puts it: “Apparently independently, H. Fleisch [(Fleisch 1947, 97–101)], R. Blachère [(Blachère 1947, 156–169)] and C. Rabin [(C. Rabin 1955, 3–4)] arrived in the forties at the conclusion that the language of the Koran, far from being pure Meccan either subsequently revised (Vollers) or slightly adapted to the poetic idiom, was none other than the poetic *koinē*.” But the seeds of this later consensus were already laid by Friedrich Schwally in *Geschichte des Korans* part II in 1919. I believe that this is the first explicit endorsement of this view, and with that also one of the more strongly argued iterations of the view. It is worth repeating the whole passage here:

Generally, any tradition connecting the ‘Uthmānic text in any way with dialectal questions must be rejected, since the Koran is not written in a local dialect at all but rather has a language identical to that of the pre-Islamic poems. These, however, cannot possibly have been written in dialectal form, as their authors belonged to quite different tribes, living so far apart that the texts would have to show strong idiomatic differences. Admittedly, when fixing a text in such a defective script as Arabic’s, where vowels are generally not indicated and many consonants are expressed by the same sign, some idiosyncrasies of the verbal presentation were simply not recognizable at all. Still, the lexical and grammatical agreement is such that an actual uniform language must be assumed. After all, given what we know about linguo-geographical conditions in other parts of the world, it would be a total contradiction if such a drastic disappearance of dialects were to have occurred in large areas of the Arabian Peninsula. We are, thus, obliged to conclude that the ancient poems, as well as the Koran, were composed in a generally intelligible standard language, the difference of which from the local dialects of cultural centres like Mecca and Medina was naturally less than from that in the more distant areas of the Peninsula.

NÖLDEKE et al. 2013, 260

I have my reservations about the confidence with which it is asserted that the pre-Islamic poems represent a linguistic unity which skips both over the strong classicizing force of the classical Arabic text tradition that without a doubt has affected the language of poetry in many ways.¹ I also believe we should raise

1 For a fascinating and stimulatingly original discussion on this topic see Foreman (forth-

doubts about the true pan-Arabian provenance of the poetry. But this is not our concern here. Rather, this book's topic is to challenge the ease with which the Quran is implicated into being part of this same intertribal pan-peninsular literary language.

Despite the view certainly predating Rabin and his cohorts in the forties, it certainly gained prominence from that point onwards. It is unfortunate that Schwally's passage is never cited, despite being a much clearer formulation of the idea than any of the authors who Rabin cites to have formulated it. It was subsequently wholeheartedly accepted, for example by Zwettler (1978, 160) who says: "most have come to agree that the *'arabīya* of the poets and the language of the Qur'ān are essentially identical and that this poetic idiom was not spoken by any group of Arabs as a vernacular tongue." Versteegh (1984, 5) follows as well: "According to the accepted opinion the language of the poems [...] became the language of the *Qur'ān* as well."

Despite the consensus that has developed on this topic, to my knowledge nobody has actually attempted to demonstrate that these two languages are "essentially identical." Instead, this has simply been asserted. Any definition of what the linguistic features of this supposed shared language are is something the field has simply remained silent on. If there is any assertion about its linguistic features at all, scholars have pointed to the Arab grammarians as having codified or standardized it (e.g. Zwettler 1978, 101, 148).

I do not wish to get into the question here whether the language of poetry and the Quran is just the 'Old Arabic' as it was spoken before Islam (e.g. Versteegh 1984) or a specific oral-formulaic register (e.g. Zwettler 1978) and therefore it seems sensible to abstract away from the terms 'Old Arabic' or 'poetic Koiné' and the many other terms that have been used by authors previously, as sometimes the same term may have a different meaning to two different authors. For example, to Rabin (1951, 3) 'Classical Arabic' is the language of pre-Islamic poetry, while Fischer (2002, 1f.) would call that pre-classical. Hence, I will define my own terms here.

First, I will use the term *ʿarabiyyah* to mean any form of Arabic that the grammarians deem fit to describe (specifically the earliest grammarians such as Sibawayh and al-Farrāʾ, but later grammarians do not seem to deviate much from them). Both Classical Arabic prose and the language attested in pre-Islamic prose (but less so Quranic prose, as we will see in chapter 3) fall

coming) who presents compelling evidence that one can in fact find systematic phonological differences between different poems, suggesting that the poetry—even after centuries of classicization in their transmission—can still display significant linguistic diversity.

within the range of variation described by the grammarians. By **Classical Arabic**, I refer to the subset of features of the *ʿarabiyyah* that eventually become a strict normative standard, this is the form of Arabic that is described in modern textbooks of Classical Arabic such as Wright (1896) and Fischer (2002). As we will see in chapter 2, this covers only a small subset of many factors of phonological and morphological variation present in the *ʿarabiyyah*. I use the term **Quranic Arabic** to refer to the register in which the Quran was initially composed. I will argue that—contrary to the opinion of other authors—the consonantal skeleton of the standard text is a fairly good guide to its linguistic features.

Of course, viewing the language of the pre-Islamic poetry and Quranic Arabic to be one and the same language is not a new view. The Classical Arabic grammarians themselves do not make a systematic distinction between ‘Arabic of Poetry’, ‘Arabic of Eloquent Arabs’ and ‘Arabic of the Quran’. All three categories belong to the same eloquent language: the *ʿarabiyyah*. As with all distinctions of language, such distinctions of course remain arbitrary. The differences between Swedish and Norwegian are small enough that there is a high degree of mutual intelligibility—yet for political reasons these languages are considered separate. On the other hand, a Berber speaker from the Souss in Morocco is unlikely to get very far conversing with a Libyan Berber from Zuara, but for pan-Berberist ideological reasons might nevertheless insist that they speak the same language.

As mentioned, while there is near-universal agreement on the idea that the Quranic Arabic and the language of poetry is the same language, the *ʿarabiyyah*, its features and supposed similarities are seldom defined. It is not of much use for a linguist to argue what should or should not be considered the “same language”. However, what a linguist can do is evaluate the linguistic features of a distinct corpus. Studying the Quran, for example, we see that it has a highly consistent use of only a *subset* of the linguistic features considered to be *ʿarabiyyah*. As already noticed by the Arab grammarians, for example, the Quran exclusively uses *dālika*, *tilka* and *hunālika* for the distal demonstratives, and never uses *dāka*, *tika/dika* or *hunāka*. In this the Quran is clearly distinct in its linguistic behaviour from Classical Arabic poetry, and Classical Arabic prose of later times, which use these forms much more freely. While there is no objective way to decide which and how many isoglosses one needs to declare something as not being “essentially identical” to the language of poetry, observing such recurrent and systematic use of only a specific subset of the linguistic variation is meaningful, and therefore important to distinguish, especially if it would otherwise lead to the uncritical acceptance that two corpora—in the case the Quran and pre-Islamic poetry—are essentially identical.

The fact that modern scholars did not sense the need to define this and demonstrate the similarities seems to stem from the fact that they assume that the language of the Quran is more or less identical to the textbook standard of Classical Arabic which is a fairly uniform linguistic system which comes to dominate later Classical literature. This, no doubt, is aided by the fact that the Quranic reading tradition most widely adhered to today, the one of Ḥafṣ Ṣan Ṣāṣim, has linguistic features that are very close (although by no means identical to) Classical Arabic. However, Ḥafṣ's tradition is just one of the two transmissions of Ṣāṣim, and besides him there are yet nine other canonical readings, each with two canonical transmissions; these readings do not just differ in the interpretation of the meaning of certain verses, but, in fact much more frequently and for our purposes more importantly, differ significantly from one another in phonetics, phonology and morphology.

Thus, we find that most modern authors assume that the *ṣarabiyyah* as reflected in the Quran *must* have had the *hamzah*, despite the fact that the grammarians describing the *ṣarabiyyah* clearly allowed for forms that had lost this, without any normative expression of disapproval (see Sībawayh 111, 541 ff.). For example, Hans Wehr in his review of Fück's *Arabiyya*, commenting on why we cannot trust the orthography of the Quran to learn about the phonology of Quranic Arabic says the following: "Die wesentlichen umgangssprachlichen Merkmale der koranischen Orthographie, das Fehlen der Nunation sowie des Stimmritzenverschlußlautes im Wortinnern und am Silbenende (*bīr*, *mūmin*, *nāyim* usw.) rühren wohl aus dieser älteren den Dialekt wiedergebenden Orthographie her und wurden übernommen, als man vor der Aufgabe stand, die 'Arabīya-Aussprache mit den Mitteln der bereits vorliegenden Orthographie zu fixieren."² (Wehr 1952, 184). The implicit assumption here is that the *ṣarabiyyah* must have had the forms *bīʔr*, *muʔmin* and *nāʔim* as in Classical Arabic as it is taught today. Moreover, it is asserted that it was the target pronunciation of Quranic Arabic. Here the overreliance on the widespread reading tradition of Ḥafṣ shows. Had Hans Wehr instead relied on the recitation of the Quran in the Maghreb, he would have indeed heard 'and a well' (Q22:45) not be recited as *wa-bīʔrin*, but instead as *wa-bīrin*, and 'A believer' (Q40:28) not as *muʔminun* but as *mūminun*, as this is how the equally valid and canonical recitation of Warṣ Ṣan Nāfīf has it, a traditions still popular and broadly

2 The essential colloquial features of the Qur'anic orthography—absence of nunation and also of the glottal stop within words and at the end of syllables (*bīr*, *mūmin*, *nāyim*, etc.)—no doubt the result from this earlier orthography which reproduced the dialects; and they were carried over when the problem arose of setting down the pronunciation of the 'arabīya by means of the orthography available at the time (translation from Zwettler 1978, 123f.).

adhered to by Muslims all over Morocco, Algeria and Tunisia.³ If, as seems to be implied here, the Quranic recitation of Warš is somehow less reflective of the *ʿarabiyyah* than is the recitation of Ḥafs, this should certainly be explained, and not be implicitly assumed.

An equally egregious example of the unintentional but nevertheless highly problematic imposition of modern prescriptive norms of Classical Arabic unto the Arabic of the Quran can be found in Zwettler's (1978, 179, n. 70) discussion of the pronunciation of *nabīʾ* for 'prophet': "the use of *hamz* in *nabīy* was a peculiarly Ḥijāzī pseudo-correction and a feature neither of the *ʿarabiyya* nor of the other dialects."⁴ Where Zwettler's certainty comes from that *nabīʾ* does not belong to the *ʿarabiyyah* is unclear to me. Not only is the form *nabīʾ* recognized in lexicographical works like *Lisān al-ʿarab* and by grammarians like Sibawayh (III, 547, 555)—never with the qualification that it is not *ʿarabiyyah*—it is once again a reading that is broadly adhered to even today in Quranic recitation in the tradition of Nāfiʿ.⁵ The only reason I can see why one would decide that this is not the *ʿarabiyyah* is due to anachronistic imposition of a much later Classical norm that in no way need be reflective of the *ʿarabiyyah* as it was conceived of at the time of the prophet—even if we would accept that the Quran was composed in that register, for which no compelling positive evidence has been presented.

Nöldeke (1910) criticized Vollers (1906), not altogether fairly, for taking the Gustav Flügel Quran as essentially the standard text, not recognizing the equal validity of reading traditions not reflected in this text edition. This criticism however should be seen as carrying a call to action: if the Quranic reading traditions do not reflect the "true" Quranic Arabic, as opposed to the Arabic of the Flügel Quran, what *do* the Quranic reading traditions represent, and

3 And he is not the only one of the canonical readers that would read these words thus. Likewise, ʾabū Ǧaʿfar and optionally ʾabū ʿamr recited like that. Also, *nāʾimūna* is read with the loss of *hamzah* by the canonical reader Ḥamzah when he pauses on it (as he likely usually would have, as both cases of it stand in verse final position, Q7:97; Q68:19).

4 One may also add that *lā tanbīr* neither means—as Rabin, cited by Zwettler, translates—"do not screech" nor as Zwettler suggests "do not raise your voice (i.e. with an expiratory stress)." *nabr* is just an alternative, and seemingly more archaic term for *hamz*. See Xalīl b. ʾaḥmad (*Kitāb al-ʿayn* s.v. نَبْر) where this meaning is given as its primary meaning (see also, Lane 2757a; *Lisān* 4323b). This obviously makes better sense in the context of a (no doubt falsified Ḥadīth) where the prophet tells his follower off for saying *yā-nabīʾa l-lāh* 'O prophete' of God', by answering *lā tanbīr bi-smī!* 'don't apply the *hamzah* to my name!'

5 Moreover, whence the certainty comes that *nabīʾ* is the pseudo-correct form and *nabīy* the proper form is also unclear. As this word clearly comes from Hebrew נָבִיָּה which, at least historically contains ʾ (as evidenced by the final ʾ), it actually stands to reason that it *should* have the *hamzah* in Classical Arabic (see § 6.5.1.1 for further discussion).

which of these (if any) *does* represent the true language? It should be clear from the previous paragraphs that many authors that have opined on the language of the Quran—although occasionally paying lip service to the existence of these readings—have essentially continued to operate on the same simplifying assumption as Vollers, assuming that their print Quran is the standard text and an accurate reflection of Quranic Arabic, and that the other 19 canonical transmissions of the Quran are just inconvenient noise to be ignored.

However, the Quranic reading traditions are not just differences on how to read a certain word or syntagm; the many distinguishing factors between the different reading traditions are in fact linguistic, marking both morphological and phonological distinctions. ʾabū Ǧaʿfar, for example, regularly loses any pre-consonantal *hamzah* (e.g. *rās* ‘head’); al-Kisāʾī has a fourth phonemic long vowel (*hadē* ‘he lead’ but *daʿā* ‘he cried out’) and Ibn Kaṭīr consistently uses long plural pronouns such as *ʾantumū* ‘you (pl.)’, *ʿalayhimū* ‘upon them’. When scholarship reached the consensus that the language of the Quran is essentially the same as that of the poetry, this should not have been uncritically accepted. One should ask which of these 20 canonical transmissions, if any, is the true language of the Quran. And why should that one be preferred over the other as being representative of the *ʿarabiyyah*? Just because the transmission of Ḥafṣ strikes the modern scholar as most familiar, as it stands closest to the strict classical standard that emerges centuries later—the standard to which also the language of poetry gets mercilessly reshaped (C. Rabin 1955, 21)—it cannot be the reason why we close the book on the investigation into what the language of the Quran is.

As should be clear from the discussion above, despite there now being well over a century of scholarship discussing the language of the Quran in relation to the language of the pre-Islamic period and the language of poetry, the Quran has never been allowed to tell its own linguistic history. Instead, its linguistic history has been co-opted by those telling the linguistic history of the poetry with the automatic assumption that the Quran is part of this same history as well. This is an oversimplification and indefensible from the perspective of historical linguistics. The Quran is a long text with ample linguistic information, not only in its reading traditions, but also in the very structure of the text itself. Scholars have gone to great lengths to disqualify the value of the Quranic orthography to inform us about its language. However, avenues that examine how one can prove that the orthography is meaningful, or indeed rightfully dismissed, have rarely been explored. The orthography of the Quran is quite distinct from the later Classical Arabic orthography, and likewise differs significantly from Arabic orthography that we find in pre-Islamic epigraphy;

this should give us pause: why is this text so different, and why was it so important to preserve?

Moreover, the Quran follows rhyming patterns throughout almost the entire text; only a tiny fraction of the verses does not rhyme with the rest. Rhymed text is an invaluable part of the toolkit of any philologist aiming to reconstruct the language of a certain period, as it helps to break through the otherwise difficult to disentangle question of what parts of an orthography are simply historical spelling and which parts reflect phonetic reality. Vollers (1906, 55 ff.) saw the value of Quranic rhyme as a source of linguistic information, but this valuable linguistic resource has been almost completely ignored since then.⁶

1.2 The Uthmanic Text Type and the Quranic Consonantal Text

One of the reasons why the Quran may not have been afforded the ability to tell its own linguistic history originates from a certain amount of skepticism that the text and its orthography, as we have it today, can be confidently projected back to the very period in which the text was composed. Criticism of the transmission history of any tradition of text is a staple of Islamic studies, and certainly in the seventies severe doubts were raised that the Quranic text truly stems from the time that the tradition tells us it comes from.⁷ However, times have changed and important advancements have been made especially when it comes to the textual history of the Quran. Early Quranic manuscripts have in recent years become easily accessible as more and more are massively digitized and editions are published. Where previously it may not have been as clear how ancient and well-preserved the Quranic text truly is, and thus what to make exactly of its orthographic idiosyncrasies, it is now clear that the text is considerably ancient. By examining the specific orthographic idiosyncrasies of the Quranic text across a variety of early Quranic manuscripts, Van Putten (2019c) has shown that all early manuscripts come from a single written archetype whose orthographic specificities have been judiciously copied

6 Notable exceptions being Nöldeke et al. (2013, 415) and Rabin (1951, 115 f., §bb) who both realized that Quranic rhyme clearly distinguishes *ā* (written with *ʔalif*) and *ē* (written with *yāʔ*) where Classical Arabic pronounces them both as *ā*. But strangely, this view was never picked up and widely accepted. Diem (1979) even goes so far as to explicitly deny that the two spellings of the Classical Arabic *ʔalif maqṣūrah* reflect any pronunciation difference at all—not addressing the evidence from the rhyme to the contrary. See Van Putten (2017a) for a defense of a phonemic distinction of *ē* and *ā* in Quranic Arabic.

7 Most notably through John Wansbrough's work, who argued that the canonical text only came to a close about two centuries after its traditional date around 650 CE (Wansbrough 1977).

over the centuries. All early manuscripts known so far descend from a single text type referred to as the **Uthmanic Text Type (UT)**,⁸ which may plausibly be connected with the standardization effort attributed to the third caliph ʿUṭmān b. ʿAffān (reign 23–35 AH/644–656 CE). While the manuscript evidence cannot preclude with absolute certainty an earlier redaction, a time much later than his reign is now a virtual impossibility. We have large portions of the Quran complete in first century manuscripts that all belong to the UT, with carbon dating early enough that much later dating is rather unlikely. In other words, we have evidence and a clear view of the Quranic text in all its linguistic details as it was written down mere decades after the death of the prophet. This should afford it the central and essential role as a primary source for the language of the Quran that it is.

The UT is highly uniform, but there are about 40 variants in the consonantal skeleton of the text in which the original different regional codices differ from one another. These regional variants can be attributed to four regions: Syria, Kufa, Basra and Medina, and these regional variants form a stemma that goes back to a single archetype (Sidky 2021; Dutton 2001; 2004; Cook 2004). These four regional types must be archetypal to the four regions and must be part of the initial distribution of the UT.

A central pillar of the current work is the incorporation of linguistic information of the UT. Whereas the consonantal skeleton of, e.g. the Hebrew bible has long been viewed as an important source of linguistic information of Biblical Hebrew (e.g. Kahle 1947, 95–102; Tov 1992, 47–49; Khan 2013, 13–30 and of course the rich Ktiv/Qere tradition within the Masoretic tradition itself), the same attention has not been accorded to the **Quranic Consonantal Text (QCT)**. As there is no critical edition of the UT, my transcriptions of the QCT will be based to a large extent on the standard text, as we find these in print Qurans today which ultimately descend from the **1924 Cairo Edition (CE)**. Unlike some other earlier print Qurans, such as the 1834 Quran of Gustav Flügel, the orthography of the CE is very conservative, and often is a fairly accurate representation of what we find in 7th century Quranic manuscripts. Due to the impressive uniformity of early manuscripts, and the fairly accurate representation of such documents in the CE, a critical version of the text is of significantly lesser importance than it would be for, say, the Hebrew Bible or forms of the New Testament. Moreover, as the CE is based on traditional medieval works that diligently docu-

8 With the exception, of course, of the spectacular find of the lower text of the Sanaa Palimpsest, which seems to represent another text tradition, plausibly from a tradition deriving from a companion codex (Sadeghi and Bergmann 2010; Sadeghi and Goudarzi 2011; Sinai 2020; Cellard 2021), for a different opinion see Hilali (2017).

mented the orthography and spelling variants of the UT, the QCT of the CE can be thought of in some ways as the result of medieval text criticism, although lacking the diligent critical apparatus that one might wish to have. Despite the impressive conservatism of the CE, the text is not always an accurate reflection of what we typically find in early manuscripts. This is especially the case for the use of *ʔalif*, which is used to write the *ā* significantly more often in the CE than is typical for early manuscripts. But there are also several other orthographic practices innovative in the CE compared to early manuscripts. For example, in early manuscripts the nominative pronoun *ḏū* is consistently spelled *ḏو*, while in early manuscripts it is consistently followed by an *ʔalif*, *ḏو*. Whenever relevant, I will change the text to better fit what early manuscripts reflect and I will reference the relevant data for such a reconstruction, often discussed in detail in Appendix B.

In my transcriptions of the QCT, I will use the consonantal dotting, even though those were used quite sparingly in early Quranic manuscripts. In fact, it is often suggested that the original redaction of ʕuṭmān's standard text had been completely without consonantal dotting, and was just a bare consonantal skeleton known as the *rasm*. Bursi (2018) has recently pointed out that there is rather little evidence for such a view. The very earliest manuscripts that we have all contain some amount of dotting—although indeed used sparingly.⁹ Moreover, he demonstrates that the sparse dotting in early Hijazi manuscripts clearly show similar patterns to the dotting used in the early documentary papyri.

The reason why I provide full dotting for the QCT, first and foremost is because it greatly aids the readability of the text. But I also believe it is justified to assume that we can know the correct dotting for the consonants for two reasons. First of all, whenever dots do appear in the early manuscripts, they overwhelmingly agree with what we find in the CE and in other manuscripts that have dots. This gives us confidence that for the majority of the text, there was a pretty clear idea what the dotting of the consonant was supposed to be, even when early manuscripts were often incomplete in marking it. The second reason is based on the evidence from the reading traditions. Sidky (forthcoming) shows that the canonical reading traditions of the Quran only disagree with each other 284 times on consonantal dotting. This may seem like a large number, but considering the thousands of times that the readers could have

9 In fact, it is rather the later Kufic manuscripts that lose (almost) all dots. One wonders if the traditional view that the UT had a bare undotted *rasm* stems from the examination of the somewhat later Kufic manuscripts, rather than the truly oldest Hijazi manuscripts with significantly more dotting.

plausibly disagreed on the dotting but do not do so, the number is actually strikingly low. This suggests that there indeed was an early consensus on what the consonants of the QCT were, even when the text may frequently have been written ambiguously. This consensus should not be overstated either, of course. Much of the dotting may be filled in by common sense, rather than a shared memory. In most contexts, there simply is no other plausible reading of the sequence *ذلك* than *ذلك*, that is *dālīka*. The combination of self-evident readings and consensus on the non-self-evident parts provide a strong basis to transcribe the QCT with these dots.

In my representation of the QCT, I do not distinguish between *ʔalif maq-ṣūrah bi-ṣūrat al-yāʔ* (e.g. *بنى* ‘he built’) and *yāʔ* (e.g. *بنى* ‘sons of (gen./acc.)’), as these are not distinct graphemes in early Quranic manuscripts.¹⁰ In non-final position, both are written with a dotted denticle (e.g. *هديه* ‘he led them’, *يهديه* ‘he leads them’).

The *hamzah* is never written in early Quranic manuscripts as it lacks any means to express it.¹¹ If the seat of the *hamzah* is *yāʔ*, it will simply be represented as any other *yāʔ* as we find in the CE, e.g. *الذئب* ‘the wolf’ and will not be displayed as a dotless *yāʔ* (الذب).¹²

The *tāʔ marbūṭah* is never distinguished from the *hāʔ*, as this graphemic distinction does not exist in early Quranic manuscripts, e.g. *نعمه الله* ‘the grace of God’.¹³ The earliest *muṣṣḥaf* that I am aware of that uses the two dots on top

10 Nor are they distinct graphemes, even in much later manuscripts, Quranic or otherwise.

11 We can hardly agree with Zwettler (1978, 179, n. 70) that the *hamzah* is most likely the oldest of the diacritical signs used in the Quranic Arabic manuscripts (citing, but misunderstanding Abbott 1939, 39f.). To the contrary, the *hamzah* sign does not get expressed with a designated sign in Quranic manuscripts for the first centuries of Islam, even when vowel signs are invented, the *hamzah* is simply expressed with vowel signs which do double duty as markers of the *hamzah*, but in that context are not always unambiguous in whether they denoted *hamzah* or just vowels. Even while at some point Kufic manuscripts start using separate colours for the *hamzah* in the third Islamic century, contemporaneous with this practice there are still many manuscripts that do not distinguish them. The modern *hamzah* sign constitutes the latest layer of Arabic diacritics (Revell 1975, 180).

12 The requirement of a dotless *yāʔ* for the seat of the *hamzah* is a result of typographic constraints, and not a practice that was commonplace in medieval manuscripts. Lack of awareness of Classical Arabic manuscript culture has occasionally led to rather bizarre statements, such as Hopkins (1984, § 20a) affirming that *بئس الرأي* “what a bad opinion!” must lack a *hamzah* because it is written with a dotted *yāʔ*. It is not difficult to find words written with a dotted *yāʔ* where the word carries a *hamzah* in a context where it is *explicit* that the word is to be read with a *hamzah*.

13 As Revell (1975, 180) points out, the dotting of the *tāʔ marbūṭah* constitutes the latest layer of Arabic diacritics, one that was only rarely introduced in Kufic manuscripts. See also Moritz *EN*, “Arabic Writing”.

of *tāʾ marbūṭah* is the Ibn al-Bawwāb Quran from 391AH/1000CE, the earliest dated Nasx Quran written centuries after the ancient Hijazi and Kufic Qurans that we are concerned with.

Whenever there is disagreement between reading traditions on how a certain consonantal skeleton is to be read, the disputed consonant will be left undotted, for example فتسوا (Q4:94) which is variously read *fa-tabayyanū* (فتبينوا) or *fa-tatabbatū* (فتثبتوا).

Finally, while in terms of the spelling of the *ʾalif* for *ā*, I will usually follow the CE, with the important note that the CE frequently spells words *plene* where manuscripts usually write it defectively, when it is relevant to the discussion, I will write the word defectively if it is consistently done so in early manuscripts, or I will use a dagger *ʾalif* to indicate that it is sometimes written with and sometimes without this *ʾalif*. Note here that it means that the dagger *ʾalif* has an explicitly *different* function in my transcription of the QCT compared to how it is used in the Quran today. It does not just note that an *ā* is written defectively, but that the *ʾalif*, which may be of any function, is not always written in the manuscript. An example might be شى to denote that *šayʾ* is variously written شى or شى in the location of the text under discussion in early manuscripts.

1.3 Overview

As it stands now, the Quran has never been allowed to tell its own story as to what its language is; instead, rather convoluted arguments have been developed as to why we cannot and should not use the Quranic orthography as a source of information for its language, all the while (implicitly) assuming that the modern standard Cairo Edition is an accurate reflection of the language of the Quran. But just as we cannot take for granted that the Quran *does* display the language as it was composed, we cannot take for granted that the Quran *does not* display the language as spoken. We *certainly* cannot assume that the language of the Quran must have been linguistically identical to the Classical Arabic or to the *ʿarabiyyah*. The current work aims to close this major lacuna in our understanding of the history of the Arabic language. I believe there are four topics that need to be addressed before we can work towards a final conclusion about the language of the Quran.

In **Chapter 2**, I will explore what the *ʿarabiyyah* is according to those that are said to have standardized it—the Arab grammarians. The *ʿarabiyyah* is all too often equated to Classical Arabic, the fairly uniform standard language described in textbooks and which forms the basis for Modern Standard Ara-

bic. But the early Arab grammarians allow for much more linguistic variation, clearly allowing and even endorsing forms that would not be considered acceptable in Classical Arabic. However, in order to understand the linguistic environment from which the Quran emerged, we must of course be informed by these earlier sources rather than anachronistically project the textbook standard onto this period and expect the Quran to conform to that.

With a clearer picture of the linguistic variation that was accepted in the *ʿarabiyyah* according to the early grammarians, it will be possible to examine the Quranic reading traditions in **Chapter 3**. As has been mentioned, the Quran today is accepted to be recited in ten different reading traditions, of which several still enjoy broad popularity. It will be shown that much of the linguistic variation described by the Arab grammarians is in fact employed in these reading traditions. Showing that the great amount of linguistic variation the grammarians describe are not just odd deviations from an implicit accepted standard language close or identical to Classical Arabic, but that this variation was inherently part of the *ʿarabiyyah* and employed as such. Moreover, the chapter will show that the linguistic variation in the reading traditions cannot be understood as dialects of Arabic. They all have clear artificial linguistic elements which must be understood as conscious decisions to change the language as part of an artificial performance register. However, due to the vastly different approaches found in each of these readings, it will be shown that it is difficult to decide what the true language of the Quran is, as the reading traditions provide twenty mutually exclusive answers.

As the reading traditions seem unable to give a unified answer as to what the language of the Quran is, in **Chapters 4 and 5** I shift focus to the language of the QCT. While some authors have admitted the possibility that the Quran in its original composition may have had some accommodation to a local pronunciation, again and again it is affirmed that grammatically it is essentially the same language as the poetic *koiné*. **Chapter 4** puts this claim to the test by comparing the isoglosses of the different Arabic dialects collected by the Arab grammarians and comparing them against the linguistic data that can be gleaned from the QCT. **Chapter 5** subsequently examines the phonological features of Quranic Arabic. While in **Chapter 3** it is shown that when doing this with the reading tradition no clear pattern appears at all, all readings haphazardly mix and match features from different dialects, when looking at the QCT a strikingly uniform picture emerges. In its morphological and phonetic features, the language of the QCT is clearly Hijazi Arabic. Occasionally this can be corroborated by pre-Islamic Arabic epigraphic evidence as well. This clear picture that emerges cannot be coincidence, and therefore we must conclude

that the QCT is an accurate reflection of the language of the Quran, and that this language is indeed the Hijazi dialect.

Finally, **Chapter 6** and **Chapter 7** return to the reading traditions. In previous articles I have shown that Quranic Arabic contained a number of features quite distinct from its later reading traditions. Most notable here is the fact that Quranic Arabic seems to have almost completely lacked the *hamzah* and had lost final short vowels and *tanwīn*. A major source of skepticism of earlier scholars was that, had the language of the Quran been so different from its reading traditions, one would expect to see traces of this in the tradition and moreover one would expect to see pseudo-correct forms. **Chapter 6** dives deep into the classicization of the *hamzah*, uncovering a variety of unetymological *hamzahs* appearing in the reading traditions as well as many places where a word should have logically had a *hamzah* but inexplicably lacks it. **Chapter 7** will focus on the many uncertainties and disagreements among the readers on the case inflection—sometimes uncovering words that unexpectedly lack case inflection completely. Moreover, these chapters focus on early reports of the readers, and discussions that followed on some of the controversial readings. These show that at least for some of the readers, for whom we have early reports, it is clear that their choices were not a matter of accurate transmission of how readers believed the prophet would have said it, but rather it was a rational endeavour that required advanced grammatical knowledge. It was the duty of the Quranic reader to rationalize and choose which words should receive which case vowel or *hamzah* and why. This clearly rationalising approach of the readers can be uncovered from the tradition.