PART I

*

LANGUAGES, WRITING SYSTEMS AND BOOK PRODUCTION

I

The languages of the Old Testament

geoffrey khan

The languages of the Old Testament are Hebrew and Aramaic. The majority of the text is in Hebrew, with Aramaic being restricted to several chapters in Daniel and Ezra,[1] a single verse in Jeremiah (10:11) and one phrase in Gen. 31:47.

Hebrew was originally a language spoken by inhabitants of Canaan. If there is a historical basis to the biblical accounts of Israelite settlement, the language must have been adopted by the Israelite tribes after their migration to the land. It is, in fact, described once in the Old Testament as 'the language of Canaan' (שְׂפַת כְּנַעַן, Isa. 19:18), though elsewhere it is referred to as 'Judaean' (יְהוּדִית).[2] The name 'Hebrew' is derived from the ancient name of the Israelites *Iḇrim* (עִבְרִים). The term is first attested as a designation of the language in the Hellenistic period in the Greek adverbial form Ἑβραϊστί 'in Hebrew' and in rabbinic Hebrew sources in the form עִבְרִית 'Hebrew'. The 'Aramaic' language is referred to in the Old Testament by the term אֲרָמִית.[3]

The earliest surviving records of Hebrew and Aramaic are inscriptions datable to the tenth century bc. Hebrew was a living language which was spoken until the end of the second century ad. Thereafter it continued to be used as a literary language until modern times. In the twentieth century a vernacular spoken form of Hebrew, based on a form of the literary language, was revived as the official language of the State of Israel. Aramaic was widely spoken in the Near East throughout the first millennium bc and the first half of the first millennium ad. Thereafter its spoken forms became geographically more restricted, but it still survives as a vernacular today in various areas.

Hebrew and Aramaic belong to the north-west branch of the Semitic family of languages. Other north-west Semitic languages include Phoenician, Moabite (known almost exclusively from the Mesha stele), Ugaritic and

---

1 Dan. 2:4–7:28, Ezra 4:8–6:18, 7:12–26.
2 E.g. 2 Kings 18:26, 28, 2 Chron. 32:18, Isa. 36:11, 13, Neh. 13:24.
3 2 Kings 18:26, Isa. 36:11, Ezra 4:7, Dan. 2:4.

GEOFFREY KHAN

Amorite (known mainly from proper names). To the Semitic family belong
also languages such as Akkadian and Eblaite, which are normally classified
as east Semitic, and a south Semitic branch that includes Arabic as well as
various languages in south Arabia and Ethiopia, including Gǝʿǝz (Ethiopic),
though Arabic is sometimes classified in a separate central Semitic branch.[4]
One of the closest relatives of Hebrew is Phoenician, which was spoken in
coastal areas of the Levant. It is attested in inscriptions in the first half of the
first millennium BC, and later in Phoenician colonies in the Mediterranean.[5]
Hebrew and Aramaic are usually classified together in a subgroup of north-
west Semitic called Canaanite, which was distinct from Aramaic. It was the
Phoenician alphabet that was used to write Hebrew and Aramaic in the early
first millennium and the scripts that were used for these languages at later
periods were all descendants of this alphabet.[6]

The Hebrew texts of the Old Testament were composed at various periods
before, during and after the Babylonian exile (597/587–538 BC), a few archaic
passages being dated by some scholars to as early as the second half of
the second millennium BC. The Aramaic passages of Daniel and Ezra were
composed in the post-exilic period. The earliest biblical manuscripts are found
among the Qumran scrolls, which date from the second century BC to the
first century AD. The printed editions that are in use today are based on a
form of text found in medieval manuscripts that derives from a school of
scholars in Tiberias known as the masoretes. The terms 'biblical Hebrew' and
'biblical Aramaic' are generally used to refer to the form of the languages that
appears in the printed editions and it is this form that is presented to students in
grammatical textbooks. The first task in describing these languages, therefore,
must be to establish the extent to which this masoretic form of the languages
corresponds to the form they had at earlier periods when the various books
of the Old Testament were composed.

The Tiberian masoretic manuscripts that have come down to us are datable
to the ninth century AD onwards. The Tiberian masoretes were active over a

---

4 For the classification of Semitic languages see A. Faber, 'Genetic Subgrouping of the Semitic Languages', in R. Hetzron (ed.), *The Semitic Languages* (London: Routledge, 1997), pp. 3–15.
5 For the Phoenician language see J. Friedrich and W. Röllig, *Phönizisch-Punische Grammatik*, 3rd edn, rev. M. G. Amadasi Guzzo with the assistance of Werner R. Mayer (Rome: Pontifical Institute, 1999) and C. Krahmalkov, *A Phoenician-Punic Grammar* (Leiden: Brill, 2001). For a survey of the main extant Phoenician inscriptions see M. G. Amadasi Guzzo, 'La langue', in V. Krings (ed.), *La civilisation phénicienne et punique. Manuel de recherche* (Leiden: Brill, 1995), pp. 19–29.
6 J. Naveh, *Early History of the Alphabet. An Introduction to West Semitic Epigraphy and Palaeography* (Jerusalem: Magnes Press, 1982); P. T. Daniels and W. Bright, *The World's Writing Systems* (Oxford University Press, 1996).

period of several centuries in the second half of the first millennium AD. Their activities ceased at the beginning of the second millennium. The components of the manuscripts deriving from the Tiberian masoretic tradition that are of greatest importance for our discussion of the biblical languages are the consonantal text and the vocalisation signs. It is important to note that in addition to these written components the masoretic tradition also contained an orally transmitted component in the form of a reading tradition, which, during the masoretic period, was passed on from one generation to the next. It is this Tiberian reading tradition that is represented in graphic form by the vocalisation signs. At the end of the masoretic period the written components of the Tiberian masoretic tradition, including the consonantal text and vocalisation signs, had become fixed and were transmitted in this fixed form by later scribes. By contrast, the oral component (i.e. the Tiberian reading tradition) was soon forgotten and appears not to have been transmitted much beyond the twelfth century AD.

The biblical scrolls from Qumran show us that during the Second Temple period a multiplicity of consonantal texts were transmitted in manuscripts. The majority of the scrolls, however, exhibit a text that is very close to the masoretic consonantal text, and have been termed 'proto-masoretic' manuscripts.[7] These differ from the medieval manuscripts only in a few orthographic details and in isolated words. It appears to have been a fixed text that had been espoused by the Jewish authorities. The tradition of the masoretic consonantal text, therefore, can be traced back to the earliest surviving Bible manuscripts in the Second Temple period. The extant proto-masoretic manuscripts show that the text had been fixed not only in content but also in orthography by the third century BC. This orthography is broadly uniform across all biblical books. It cannot, however, have been the original orthography of all the books that was used when they were first committed to writing, since inscriptions show us that in earlier centuries in the pre-exilic period the orthography was more defective, with vowel letters used more rarely. Hebrew orthography gradually employed more vowel letters as time progressed. At some stage an attempt was made to impose a standard orthography on the entire text. By comparison with independently attested epigraphic material, scholars have dated the broad profile of the orthographic practices fixed in the proto-masoretic text approximately to the period 500–300 BC.[8]

---

7 Cf. Tov, *Textual Criticism of the Hebrew Bible*, pp. 24–79.
8 Cf. Andersen and Forbes, *Spelling in the Hebrew Bible*; D. N. Freedman, A. D. Forbes and F. I. Andersen, *Studies in Hebrew and Aramaic Orthography* (Winona Lake, IN: Eisenbrauns, 1992), p. 15.

GEOFFREY KHAN

The components of the biblical text that are datable to the pre-exilic period
are of diverse origin, with regard to both time and place of composition. This
diversity is reflected in linguistic differences, especially between the archaic
sections of a poetic nature and the prose sections. In the Second Temple period,
when the orthography of the written form of the earlier texts was updated and
standardised, the language of the texts was in principle not updated, at least
not in any systematic and radical way, otherwise the aforementioned diversity
would have been eliminated. There are, however, some cases where linguistic
adaptation appears to have taken place, mainly in archaic poetic passages that
were no longer understood in the Second Temple period. In general it can
be said that the editorial activity relating to the linguistic structure of the
written text had the purpose of expressing the current interpretation of the
received form of the language rather than undertaking a linguistic reform.
The orthographic adaptation itself was not a systematic *replacement* of the
earlier orthography so much as an *expansion* of the latter by the addition of
vowel letters to reflect the current way in which the text was interpreted and
read. Although there appear to have been some scribal errors in transmission,
the core of the earlier orthography was retained. This editorial activity in
the early Second Temple period was associated also with the shift from the
Palaeo-Hebrew script to the square script, which was adopted from Aramaic.

Some of the later biblical books were composed in the period when the
fixing of the orthography took place. Their orthography, therefore, should,
in principle, be regarded as reflecting the usage that was current during the
time of their composition. In fact, the orthography of the later books exhibits
a slightly greater tendency to use vowel letters than does that of the pre-exilic
books; for instance the name of King David is spelt defectively in the pre-exilic
books (דָּוִד) but with the vowel letter *yodh* in some of the later books (דָּוִיד). The
scribes of the later books apparently aimed at adopting the by now standardised
type of orthography but were influenced to some extent by a slightly more
advanced type of orthography that developed in the Second Temple period.

The orthography of the Qumran biblical scrolls demonstrates that there was
not only diversity in the types of biblical text but also diversity in the way the
text was read. These diverse types of reading reflect differences in phonology
and morphology. A similar linguistic diversity is exhibited in the non-biblical
texts from Qumran. At this period Hebrew was still a vernacular language and
it is likely that the background of much of the aforementioned diversity in the
manuscripts reflects dialectal differences in the vernacular. The standardised
orthography found in the proto-masoretic manuscripts, which formed the

basis for the orthography of the masoretic text, reflect the reading of the Hebrew with a particular pronunciation and set of morphological forms. The Qumran manuscripts show us that this type of phonology and morphology was only one of several varieties that existed in the Second Temple period. One form is not necessarily more ancient than another. A further level of linguistic diversity can be identified if we take into consideration the reading tradition reflected by the Tiberian vocalisation. We must now, therefore, examine the linguistic background of the Tiberian vocalisation to determine how to assess this phenomenon.

The Tiberian vocalisation consists of a set of signs that were written below, above and sometimes within the letters of the consonantal text. The vocalisation system includes signs to represent vowels and also signs to represent syllable division (*shewa*), consonant gemination, the distinction between the two types of pronunciation of the so-called *bgadkfat* consonants (*dagesh*) and the consonantal pronunciation of a letter (*mappiq*). The vocalisation notation, in fact, marks more than phonology. It reflects syntactic divisions, in that it marks differences between the pronunciation of words that occur at syntactic pauses and those that occur within syntactic units. The *dagesh* sign is sometimes used, moreover, to distinguish meaning. A few isolated cases of this are found in the Tiberian tradition; the *dagesh* is used, for instance, in the *lamedh* of the word לֹא when collocated with the homophonous word לוֹ (e.g. Prov. 26:17 לֹא־לְוֹ).[9]

The vocalisation signs are a written notation that was developed by the masoretes to record a reading tradition. In the time of the Tiberian masoretes, and also for a certain period after their activities ceased, both the Tiberian sign system and the Tiberian reading tradition were regarded as authoritative.[10]

9 Yeivin, *Introduction to the Tiberian Masorah*, pp. 49, 294. One may perhaps identify this marking of *dagesh* to express a semantic distinction in its occurrence in the prefixes of imperfect consecutive verb forms to distinguish them from imperfect forms with conjunctive *waw*. Its usage is more frequent in manuscripts with Babylonian vocalisation; cf. Yeivin, *The Hebrew Language Tradition as Reflected in the Babylonian Vocalization* [Hebrew], pp. 355–63. This is not necessarily an artificial linguistic phenomenon. One may compare the use of emphasis (velarisation) to distinguish homophonous words of different meaning in Neo-Aramaic dialects, e.g. in the Christian Barwar dialect: *bera* 'well, cistern' and *beṛa* 'light'.

10 Some of the masoretes were closely associated with the Jewish authorities, e.g. Pinḥas Rosh ha-Yeshiva ('head of the academy'), who lived in the ninth century. The 'academy' (*yeshiva*) was the central body of Jewish communal authority in Palestine; cf. M. Gil, *A History of Palestine 634–1099* (Cambridge University Press, 1992), pp. 495–501.

Various other vocalisation systems existed in the Middle Ages. These include the Babylonian and Palestinian systems, which, although reflecting different pronunciation traditions, exhibit various degrees of assimilation to the Tiberian system in the extant manuscripts. The Tiberian vocalisation system soon became the standard one and replaced all other systems in the transmission of the Bible. The transmission of the Tiberian reading tradition, on the other hand, soon came to an end. As a result, the Tiberian vocalisation signs came to be read according to the various local traditions of Hebrew pronunciation. It is only recently, by studying previously neglected medieval sources, that the original Tiberian reading tradition has been reconstructed. This differs from the descriptions that are found in modern textbooks of biblical Hebrew, all of which present a form of pronunciation that was not that of the Tiberian masoretes.[11]

In a large number of places the reading tradition (*qere*) that is reflected by the vocalisation does not correspond to the consonantal text (*ketib*). Some elements of the consonantal text are regularly read in a way that does not correspond to what is written. This applies to the reading of some elements of morphology, such as the pronominal suffixes. The second person masculine singular pronominal suffix, for example, is written ךָ- but read – khā̆, without a final vowel. The verbal inflectional suffix of the second person masculine singular is written תָ- without a final vowel letter but is read – tā̆ with a final vowel. The third person masculine singular pronominal suffix on plural nouns is written יו- with a medial *yodh*, presumably reflecting a pronunciation such as -*ew*, but is read – ā̆w without the medial *yodh*.

The most satisfactory explanation for this phenomenon is that the reading was a separate layer of tradition that was closely related to, but nevertheless independent from, the tradition of the consonantal text.[12] Contrary to a view that is still widely held today, the reading tradition was not a medieval creation

11 Some examples of features of the Tiberian reading tradition that differ from the descriptions in the existing grammatical textbooks are the following: (i) the vocalic *shewa* was pronounced in most contexts as a short /*a*/; (ii) the vowel *qames* had a back rounded quality /*å*/ both when long and when short, e.g. חָכְמָה [håḵmå̆] 'wisdom'; (iii) the vowels *patah* and *segol* were long when stressed or in unstressed open syllables; (iv) the consonant *resh* was pronounced in two ways, one alveolar and one uvular. For a description of Tiberian pronunciation see Khan, 'The Tiberian Pronunciation Tradition', and G. Khan, 'Tiberian Hebrew Phonology', in A. S. Kaye (ed.), *Phonologies of Asia and Africa*, vol. 1 (Winona Lake, IN: Eisenbrauns, 1977), pp. 85–102.
12 I take the view here of scholars who have stressed the oral dimension of the text reflected by the vocalisation; cf. especially J. Barr, 'A New Look at the Kethib-Qere', *Oudtestamentische Studiën* 21 (1981), 19–37; J. Barr, *Comparative Philology and the Text of the Old Testament* (Oxford University Press, 1968), pp. 194–222; Morag, 'On the Historical Validity of the Vocalization of the Hebrew Bible', 307–15.

of the masoretes but was an ancient tradition that the masoretes recorded by their notation system. There is no evidence that the masoretes reformed the reading tradition and felt free to introduce linguistic innovations of their own.[13] The morphological features of the Tiberian reading tradition that differ from what is represented by the consonantal text are reflected already by Qumran manuscripts in the Second Temple period, for instance the second person masculine singular suffixes כה-, תה-, and the third person masculine singular suffix on plural nouns ו- without *yodh* in manuscripts exhibiting full orthography.[14]

As we have seen, in the Middle Ages various ways of pronouncing biblical Hebrew are reflected in different systems of vocalisation. The Tiberian, Babylonian and Palestinian systems of vocalisation not only use different sets of signs but also reflect clearly distinct forms of pronunciation. In addition to these three traditions of pronunciation, there is the Samaritan tradition, which was not recorded in written notation but has been passed down orally. Although the Tiberian, Babylonian and Palestinian systems differ from one another, it is clear that they are closely related in comparison to the Samaritan pronunciation of Hebrew, which is significantly different from all three. We can identify two broad streams of pronunciation tradition, the Samaritan and the non-Samaritan. The close relationship of the Babylonian reading tradition with the Tiberian and Palestinian could be explained as a result of its being transferred from Palestine to Babylonia by Jewish scholars after the Bar Kokhba revolt. A number of the differences within the non-Samaritan group appear to have arisen by convergence with the vernacular languages. This applies especially to the Palestinian pronunciation, which exhibits many features that are characteristic of Aramaic, the vernacular of the Jews for most of the first millennium AD.[15] The Tiberian system appears to have been very conservative and was largely unaffected by vernacular influence. In the Middle Ages the Tiberian reading tradition was the preserve of a small number of scholars who had received special training. The Palestinian pronunciation, which was close to the Aramaic vernacular, was far more widespread. The Sephardi pronunciation traditions of Hebrew, which are still followed today

---

13 The view that the masoretes were language reformers was held by P. Kahle, see his book, *The Cairo Geniza*. His arguments were convincingly rebutted by E. Y. Kutscher, 'Contemporary Studies in North-Western Semitic', *Journal of Semitic Studies* 10 (1965), 21–51 and J. Barr, *Comparative Philology and the Text of the Old Testament*, pp. 214–17.
14 Kutscher, *The Language and Linguistic Background of the Isaiah Scroll*, pp. 442–3.
15 The vowel system of some forms of Palestinian Hebrew pronunciation, for instance, seems to be very close to that of Jewish Palestinian Aramaic. Texts with Palestinian vocalisation also exhibit a number of features of Aramaic morphology.

geoffrey khan

in many of the eastern Jewish communities, are derived historically from Palestinian pronunciation. The Babylonian pronunciation, which was also more widespread in the medieval Jewish communities than the Tiberian pronunciation, has survived down to the present day in the reading traditions of the Yemenite Jews.

It is not possible to demonstrate the historical depth of Tiberian phonology as a whole. There is evidence, however, for the deep historical roots of certain features. One example that demonstrates the conservative nature of the phonology is the pronunciation of the *pe* in the word אַפַּדְנוֹ 'his palace' (Dan. 11:45). According to medieval sources this was pronounced as an emphatic unaspirated stop, whereas the letter *pe* with *dagesh* in all other places in the reading tradition was pronounced as an aspirated stop. The hard pronunciation of the *pe* is also mentioned by Jerome, who states that it is the only 'Latin' *p* in the entire Bible (*p* in Latin was regularly pronounced as an unaspirated stop). The word is in origin a loan from Old Persian. The unaspirated pronunciation of the *pe*, which is uncharacteristic of Hebrew, evidently preserves a feature that existed in the pronunciation of the source language.[16] The fact that this feature, which conflicted with normal Hebrew pronunciation, should have been preserved from the original period of composition right down to the period of the masoretes, centuries after contact of the transmitters of the tradition with the source language had ceased, demonstrates the remarkable conservatism of the Tiberian reading tradition. This feature of pronunciation was lost to knowledge after the Tiberian reading tradition fell into oblivion in the later Middle Ages, and it does not appear in modern textbooks of biblical Hebrew.

The lack of correspondence of some forms of pronunciation attested in the Second Temple period with the Tiberian reading tradition should not lead us to conclude that the Tiberian tradition is necessarily of a chronologically later origin. It is likely that a form of pronunciation that is very close to the Tiberian tradition existed in Second Temple times alongside other traditions of pronunciation. The Septuagint, datable to the Second Temple period, contains transcriptions of Hebrew words, mainly proper names, which appear to reflect a pronunciation that is more archaic than that of the Tiberian tradition. These transcriptions, for example, often have an /a/ vowel in an unstressed closed syllable (e.g. Μαριαμ) where in Tiberian Hebrew it has developed into an /i/ (מִרְיָם). This, however, need not be interpreted as demonstrating the

16 Cf. R. Steiner, 'Emphatic פ in the Massoretic Pronunciation of אַפַּדְנוֹ (Dan 11:45)' [Hebrew], in H. Ben-Shammai (ed.), *Hebrew and Arabic Studies in Honour of Joshua Blau* (Tel Aviv and Jerusalem: Hebrew University, 1993), pp. 551–61.