

CHAPTER ONE

ALPHABET SOUP

1. ROADMAP

This dissertation explores the process undertaken by medieval writers to produce Portuguese-language texts using the letters of the Hebrew alphabet. Through detailed philological analyses of individual texts, I focus on the strategies for using one set of characters to write a language in the context of more traditional adaptations, i.e. the conventional orthography. I examine the linguistic features of this corpus in order to challenge the conception of its writing system as marked or marginal, a view that misleadingly identifies a language with a script and with the conventions applied to that script in representing the language. Thus in contrast to past work on this topic, I use the term *adaptation* rather than *conversion*, since the latter suggests an inherent link between the source script and written language, and implies that this script is used in a derivative manner. I argue that the adaptation of Hebrew script for this Ibero-Romance language is not derivative of or secondary to Roman-letter writing, nor entirely dependent upon the conventions of written Hebrew itself. I further argue that it should not be viewed as a process performed anew by each writer and influenced primarily by spoken language (that is, as a transcription). The perspective I adopt, then, rejects the premise that the patterns manifested in this unconventional writing system are *ad hoc* creations by its writers, that the writing system requires extra effort from its readers, or that it is a less "native" system than the dominant, more conventionalized, Roman-based adaptation

that normally bears the title "written Portuguese." It also rejects the view that that the language articulated in this writing system has an intrinsic Judaic character beyond the script itself (though it does not exclude this possibility). I argue instead that writing Old Portuguese in Hebrew script was no more problematic an act than any written gesture in which multiple influences and competing conventions must be negotiated.

In Chapter 2, I present a case-by-case survey of the adaptation of Hebrew script for languages other than Hebrew, leading to a more detailed analysis of the Judeo-Portuguese writing system in Chapter 3. The three subsequent chapters offer critical editions and philological commentary on texts representing the bulk of the corpus of Old Portuguese in Hebrew script. In Chapter 4 I present a new edition of the only text to be published in edited form prior to my own work, a manual for manuscript illumination entitled *O livro de como se fazem as kores*. Chapter 5 presents a portion of the largest manuscript in the corpus, an astrological treatise known as *O livro de maḡika*. Chapter 6 presents editions of three shorter Portuguese texts written in Hebrew script, two prayer books that contain vernacular instructions for the Passover meal along with a short medical prescription. In chapter 7 I summarize the main contributions of this corpus to Portuguese historical linguistics in terms of archaism and vernacularism in the lexicon. And in the final chapter, I offer a proposal for a Judeo-Portuguese "alphabet," along with a sketch of some further problems of adaptation and interpretation that arise from the misconception of "native" versus "foreign" scripts and from the process of making Hebraicized texts "accessible" (and the preconceptions of audience and purpose on which this notion is based), as well as the adaptations that generate Hebrew characters from Roman-letter keyboards.

2. WRITING'S "SECOND MOVEMENT"

According to the best estimates, spoken language emerged about 200,000 years ago (Fischer 2001); at present there are approximately 6,000 spoken languages in the world, in addition to an unknown number of once-spoken, now-defunct languages. By contrast, there have perhaps been only a few dozen historically-distinct sets of graphic characters that have made possible the activity we call "writing." The oldest extant writing, in fact, dates from only a few thousand years ago. For two aspects of language so intimately linked, speech and writing have remarkably little history in common.

As an autonomous technological achievement, writing seems to have been "invented" only a few times in history. It first appeared five thousand years ago in Mesopotamia, while its most recent appearance as independent innovation occurred two thousand years ago in Mesoamerica. And therein lies an incongruity: for all the writing systems that exist in the world today, the *idea* to write in a society with no prior exposure to written language was implemented only a handful of times, if that – it is arguable, as Fischer (2001) puts it, that no one has ever independently "re-invented" writing (though perhaps the Mayans, separated from Eurasia by two oceans, can claim to have done so). Every subsequent instance of a language made readable for the first time has been undertaken by someone familiar with the most basic principles and practices of writing, someone capable of implementing these principle and practices in order to write another language that had to that point been invisible.

In this study I will not be concerned with the "principles and practices" devised by the very first human writers or those who produced the first written language in their cultures – however greater an achievement their technological innovation may ultimately be. This "first movement" in the history of written language is well-trodden terrain in terms of the data (though ever-more ancient writing may yet be discovered) and the theories, from ancient times through the Middle Ages and the earliest days of comparative philology to recent standard-bearers such as Gelb (1952). It is also very much active terrain, with many English-language book-length studies on writing systems and the history of writing currently in print, ranging from coffee-table art-history surveys to more linguistically-oriented studies,¹ both striving to encompass any and all forms of visible language.

What I am concerned with instead is a "second movement" in the evolution of writing. In contrast to the first innovation, this movement has been a more recurrent one, a scenario that has played out any time a writer has attempted to represent elements of his language in a novel graphical form – but, crucially, with the initial innovation of writing already known to that writer or his community. Roughly speaking, it involves taking the units of a writing system (be they the concrete entities we call letters or other more abstract conventions) that have been put to use for one language and applying them in the service another language: in other words, an *adaptation of script*. In a more restricted context, this is what the earnest *sofer stam* – not to mention the hapless French bureaucrats – were attempting to do using their respective writing systems when confronted with the sound-sequence of my name.

¹ These are too numerous to list, but two recent publications that represent each format respectively and that also bear suitably emblematic titles are Christin (2002) and Rogers (2005).

Despite its recurrence in both the history and current practice of writing, however, this "second movement" has received surprisingly little attention in both popular and scholarly writing. In the following sections I outline the different forms that this phenomenon may take, in order to single out those that will be relevant for the present discussion of script adaptation.

3. BORROWING, WRIT LARGE

Whenever a writer sets out to compose original written material in her language, be it 3,000 years ago or last Tuesday, there is the possibility that she will need to make readable to her audience a word or phrase that she knows in some way comes from beyond her language. Although the writer may know that the word didn't first appear in the world as part of her language, for her present purposes it does now belong, and so in principle can be written alongside the native material. At a particular moment in time, or over the course of many years, one or more forms for representing this word may become conventional in the orthography that the writer has learned. In order to arrive at this stage, however, previous writers would have experimented or improvised with different ways to spell this foreign material before subsequent writers (and readers) came to regard one or more of these forms as belonging to their language. At its most fundamental, this is the cognitive and mechanical act that constitutes the focus of this study: what has been or may be tried graphically to make linguistic material that has been perceived as foreign readable within another linguistic framework; what factors might enable one form or another to be favored by an individual writer and perhaps "win out" over time in the orthography; and what are the causes and consequences, linguistic or otherwise, of these orthographic gestures.

3.1. *Loanwords*

Distracting from this facet of script adaptation, however, are the inert or conventionalized loanwords. Like the bulk of the written language, these are orthographic fossils that – notwithstanding their current utility – may convey information (apart from semantic content) from a previous time, under a different set of conventions and influences. Consider a simple homegrown example: while the orthographic forms <czar> and <tsar> may tell us something about the interaction of Slavic and English spelling conventions that competed in some writers' minds at some time in the past, the decision for a literate writer today to use one or the other of these forms is at best an arbitrary choice between two established variants (unless, of course, the writer has some non-linguistic incentive for favoring one or the other: political motivations, total ignorance one of the variants, etc.). In any event, this type of alternation is not my present concern.

3.2. *Ad-hoc nativization*

The scenario shifts, however, when the writer is dealing with the task of putting foreign material into written form for the first time (as far as the writer knows). I would exclude from this consideration the writing practices of emergent literates, such as children, or adults learning to write for the first time – people whose experience with written language is not (yet) complete. My intention is to focus on the experience of literate writers trying to spell something they have never seen spelled before. It is here that loanwords become dynamic objects and that the "principles and practices" of the writing system actually do come into play. So when the students in my French course

must write down the name of a Senegalese dish that they hear for the first time from a Camerounian French speaker, they are very much required to concoct a spelling for this word by actively extrapolating the conventions of Roman script at their command (as shaped by English orthography, though in some cases conditioned by their incomplete command of French spelling patterns). This process is not restricted to isolated words: when I want to write e-mail to a friend from a Yiddish summer course, I must devise, more or less on-the-fly, a way to use Roman letters to write the continuous prose of my message in a language that I had learned using the Hebrew alphabet.

Admittedly, this last scenario is somewhat disingenuous, given that I am familiar with prior conventions in both academic and non-academic circles for writing Yiddish in Roman script. Yet what it illustrates remains, namely the writing that is produced when a particular writer abandons the conventional orthography and fuses together another set of influences. For a more illustrative example, consider the transcriptions used in a typical language-learning manual. Here the conventions of the learner's writing system may be imposed on the script to represent the spoken forms of the target language in a way that, while not necessarily on-the-fly, certainly pays little heed to whether the forms look conventional in the orthography of either the learner's native language or his target. Indeed, this is *ipso facto* the case if the languages are normally written in different scripts. The "language handbook" setting in fact represents an overlooked locus of borrowing on a much grander scale, where it is not individual items that are borrowed, but rather the graphic sheath itself.

3.3. *Borrowed scripts*

On that note, recall the distinction made earlier regarding the first and second "movements" in the production of written language, between its initial invention and any subsequent innovations. Thus far I have mostly been concerned with borrowing in the usual sense, i.e. with the incorporation, however fleeting, of elements from one (or more) language(s) into the conventionalized writing system of another. As a heuristic I will refer to this "receiving" system as the *matrix*, and to the incorporated elements as the *target*. In this scenario, then, it is the target elements that are adapted to fit a matrix defined by a writer's linguistic history and socio-political context. In such cases the matrix and the language of composition are not usually distinguished, even if the script is acknowledged to have been imported at some time in the past. It is thus conventional to talk about, for example, "the English writing system," as it would similarly be to refer to the incorporation of target elements (i.e. borrowings) into an "English matrix."

Now recall that figure about a few dozen graphic systems in the service of several thousand human languages. Naturally not every one of those languages has (yet) appeared in writing, so the real disparity may not be quite as shocking. Yet without much reflection we may take it for granted that in order to get target elements into writing at all (whether their language of origin had previously appeared in writing or not), the graphic matrix must be adapted and grafted onto those target elements. Or, viewed from the opposite perspective, it is the relatively flexible target elements that must be adapted to fit the mould of a relatively rigid graphic matrix.

What is crucial to recognize is that in such cases we are no longer dealing with a text written in the language conventionally identified with the

matrix. That is to say, the language that is being written is the target itself, rather than just items selected from it and inserted into a predefined matrix. Adapting the terminology of the only English-language book-length study on the topic (Wellisch 1978), the matrix has been *converted* in order to write the target. The term is somewhat of a misnomer, however, since the basic graphic identities of the units in the adopted script are rarely modified (in the act of "conversion," that is – the forms of graphemes may change for other reasons). Instead these units are be *adapted* – perhaps augmented with diacritic markings to capture distinctions that they do not in their conventional context – or else they are simply not selected for use at all. The written product resulting from this adaptation will certainly depend on whether the target was previously unwritten (e.g. in adapting the Roman alphabet to write a Native American language) or written in some other script (e.g. in adapting the Roman alphabet to write Turkish) – to say nothing of the idiosyncrasies of the individual adapter at work. But in either case the essential idea is the same: the graphic product employing the matrix script is not a text written in the matrix language, and may in fact bear no immediate linguistic relation to it. To drive this point home: there is nothing intrinsically Latin about the text on this page, for which I am using – in order to write prose in a Germanic target – a script that is indigenous (more or less) to an Italic matrix.

3.4. *Adapted scripts: permanent vs. transient*

There is a further distinction in this typology of script adaptation that applies more globally and socio-politically, albeit to only a handful of scripts. Script adaptation as defined above has historically been the consequence of some political event. After an invasion or political domination of some kind,

the language of the newly-dominated population usually comes to be written (assuming it is maintained) in the script used by the invaders (assuming they belong to a literate society). Thus for a population that converts (or is converted) to Christianity, the indigenous language will sooner or later come to be written in the alphabet of Christianity, namely the Roman script. The same applies to Arabic script when a population becomes Muslim or otherwise dominated by Islam, or to Cyrillic when a population is made to follow the Eastern Orthodox church. This last case is doubly interesting, insofar as Cyrillic script has further spread through Russo-centric but anti-religious Soviet dominance.

Although the Romanization, Arabicization and Cyrillicization might account for the origins of many writing systems – if not a sizeable proportion of the gross writing generated in the world – an adaptation of script need not be permanent. For instance, in separating the religious from the political in the public sphere, the Turkish language went from using Arabic script (which had itself been adapted at some point in the past under Islamic influence) to Roman script. Even more recently, some of the former Soviet republics, which just sixty or seventy years ago had Cyrillicized their writing systems, have since de-Cyrillicized them. And in the most intriguing cases, ethnic/religious factions in the community have agreed to disagree: Serbo-Croatian writers use either the Cyrillic or Roman script depending on their religious affiliation, and writers of Hindi or Urdu will use the Devanagari or Arabic script respectively for what is arguably still a single language.² Given that the history of so many

² One could mention here the various Romanization movements around the world, whose goals are to abandon a script adopted at some point in the past and elaborated over time in favour of an adaptation of Roman script, which proponents usually tout as more "efficient" in some way. This efficiency is really little more than utility, however, a deference to the political and economic spread of Roman script.

written languages is marked by the adoption and adaptation of more than one script, it is clear that the compatibility of a language and the script(s) used to write it is not based on linguistic criteria alone. It is largely for non-linguistic reasons that one script (and a set of conventions associated with it) may become so closely identified with a language that their compatibility, however fraught with difficulty, is not normally questioned by readers and writers.

Serbo-Croatian and Hindi/Urdu may be among the few modern cases of unrelated scripts in (semi-)peaceful co-existence; yet the circumstances that led to this situation must have been more common in the past, before the modern-day march of nationalization and standardized written languages. Consider in particular the situation in medieval Spain, where Ibero-Romance languages were written in one of three scripts (Roman, Arabic, or Hebrew), again depending on the writer's religious affiliation. The Hebrew alphabet in particular was used by Jews to write in Hebrew, Arabic, and Spanish, including some of the earliest-attested material in Spanish. This situation, in which Hebrew script was used to write both the vernacular in addition to the non-native Hebrew language, was far from uncommon in medieval Jewish communities. The circumstances of medieval Spain would also apply to the Jewish community (or communities) in adjacent Portugal, though records of the practice there are not as abundant as those from Spain. Moreover, the lusophone Jews who wrote in Portuguese after their expulsion in 1496 did so, based on the extant record, using Roman script exclusively. The texts that form the basis of this study – medieval Portuguese written in Hebrew script – thus represent valuable attestations of a poorly-documented but culturally-significant practice.

What makes these documents of such high interest and importance is their utter markedness: because the Hebrew alphabet was not and is not the usual one used by writers of Portuguese, the modern reader must abandon the normally-unchallenged assumption that "written Portuguese" implies "Portuguese written in Roman script." And yet it is highly unlikely that this sort of mental exercise was required of the medieval reader. The very existence of the texts, as well as the occurrence of the practice in other linguistic settings, presupposes an audience with the appropriate literacy. After all, though the matrix may be Hebrew, the target is Portuguese.

3.5. Hebrew script for non-Hebrew writing

Like the other scripts associated with major religions, Hebrew has at various times been used by Jews to write a set of unrelated spoken languages, whether within the matrix of Hebrew-language writing or in a wholesale adaptation to a target language. Arabic, Greek, most of the Romance languages, German, Persian, Turkish, and others have all been "Hebraicized" by Jewish writers (these are briefly profiled in the next chapter). Yet the Hebraic adaptation stands apart from the other major movements in two crucial ways: (1) even in the earliest adaptation of what was originally a local variant of an imperial Aramaic script, Hebraicization has never been associated with political power, and (2) a Hebraicized form of writing has persisted in only one modern orthography, namely Yiddish³ – no other adaptation of Hebrew script has persisted into the twenty-first century. These two features make the use of Hebrew script for languages other than Hebrew,

³ Although it boasts a half-millennium of texts written in Hebrew script, Judeo-Spanish writing has been largely re-Romanized since the mid-twentieth century (cf. chapter 2 § 3.2.1).

and the detailed analysis of any one of its manifestations, a unique laboratory for the study of script adaptation.

Scholars such as Goerwitz (1996: 487) have noted the distinctive place and import of Hebrew in the history of script adaptation: "The story of the Jewish scripts is ... a great deal more than the story of sectarian orthographic tradition: It is an important chapter in the history of writing." In fact, language historians get rather excited by texts whose marked or unconventional script might yield information that might not emerge from material written in a more tradition-laden one, as Paper (1977: 103) explains in reference to Hebraicized Persian texts:

It is an interesting and fascinating fact that the half-dozen earliest bits of direct textual evidence that testify to the appearance of Modern Persian are texts written ... in the Hebrew alphabet! We have, in essence, a practically continuous chain of direct linguistic data sources for the Modern Persian language in this particular Hebrew orthographic garb.

These early Hebrew-letter texts stand out because what became the conventional mode for writing Persian is an adaptation of *Arabic* script. Historians of Persian would no doubt be equally excited by Arabic-letter texts containing so-called *vernacular spellings* – forms that flout the conventional orthography by attempting to map the sounds of the language to the symbols of the script as unambiguously as possible, without regard for the usual cross-dialectal utility of standard orthography. Language historians believe that vernacular spellings are more likely to occur in early attestations of a language, making such documents more transparent with respect to pronunciation. Yet they often face a dearth of vernacular texts in the conventional script, so that attestations of the language in another script are

valued as more phonetically transparent than vernacular spellings found in the unadapted "native" script. This position is implicit, for example, in de Lange's discussion of Hebraicized Greek texts:

There is a shortage of vernacular [medieval] Greek texts that endows these fragments [from the Cairo Genizah] with particular value, but even greater interest attaches to the fact that they use a non-Greek alphabet, *and therefore convey the pronunciation* [my emphasis]. It is my impression that many of the words and some of the forms are otherwise unattested (1996: iii).

Since it is assumed that there is a lack of convention for associating the sounds of the language with the symbols of this "foreign" script, the writer's motive in what is construed as an inherently *ad hoc* process must be to spell what he hears (or what he believes should have been heard). Yet it would be misguided to characterize de Lange as claiming that the non-Greek script was used as a phonetic transcription of the language at that time; such a system is not likely to have served the needs of skilled readers, nor is it likely to have been produced by skilled writers (Wright 1982). Thus it cannot be assumed that every orthographic variant corresponds to a phonological one. Similarly, it has been claimed, particularly in reference to Judeo-Romance writing, that any Hebrew words found in Hebraicized writing must be considered unassimilated in the spoken language because scribes do not alter their spelling to conform with the system used for the vernacular part (Freedman 1972). Yet skilled readers of many languages are accustomed to dealing with multiple or competing conventions, and they usually pronounce such words in perfect harmony with the orthographically-assimilated ones. That is to say, unconventionality is not equivalent to absence of convention.

In his study, Wellisch (1978) devotes a short section to "Hebraification," and notes the unique aspects of the Hebraic tradition:

[T]he Hebrew script, the means of graphic expression for one of the smallest ethnic groups in the Roman Empire, survived not only the scripts of other more numerous and powerful nations but also the empire itself, and it attained a status on a par with the two alphabets that had also developed from the North Semitic one (p.120).

Hebrew is probably the only script that has been used to write languages other than the one for which it was originally devised without any coercive system of a religiopolitical nature (p.121).

These are indeed apt observations. What he appears to mischaracterize, however, is the genesis of this tradition:

[It was used]... initially to record foreign names (for which there already existed a long-standing tradition in the Talmud for rendering Greek, Latin, Egyptian, and other names in Hebrew transcription). This practice was later extended to words and short phrases in other languages until it was quite natural to write the vernacular entirely in Hebrew characters (p.121).

On the one hand, this description creates the impression that as Hebrew texts began to fill with loanwords, they simply morphed into another language – a gradual relexification of sorts. Yet as I described it above, the adaptation of a script is a discrete, not continuous, process. Thus, for example, no matter how many French words Chaucer and other Middle English authors inserted into their writing, their language of composition was never French (Thomason and Kaufman 1988). If, on the other hand, Wellisch's description implies that by inserting vernacular loanwords into Hebrew texts, Jewish writers could

extend the principles they devised for writing these individual items to longer passages in other vernaculars, it is difficult to account in general for many of the widely-varying patterns that characterize the earliest Judeo-Romance writing, and in particular for the (apparent) absence of "words and short phrases" in Portuguese within Hebrew texts prior to the emergence of a fully-Hebraicized Portuguese writing system.

Wellisch's description also mischaracterizes what was "natural" about extending the use of Hebrew script to other languages. It was not the orthographic dexterity they developed that persuaded Jews to write in Hebrew script, but rather the very nature of writing and (religious) identity in medieval Europe. The corpus of Old Portuguese in Hebrew script confronts the researcher with a well-elaborated product of script adaptation, carried out without coercion, but which was not maintained into modern times. This is no doubt related to Hebrew's lack of association with political sovereignty in general, and the fate of Portuguese Jews in particular. And yet there remains a writing system used by highly-skilled writers and digested by highly-skilled readers. It is almost inconceivable, in fact, that these writers and readers were not multi-literate, given how un-Hebrew-like this writing system usually is. Furthermore, the unconventional alphabet is the only markedly Jewish feature of much of the corpus – a feature that only serves to illustrate how Portuguese Jews were integrated members of their linguistic community. This Portuguese corpus shows none of the marginality and parochialism that tends to be associated with some Jewish languages (or at least some studies of Jewish languages). What it does express is an interest in and a market for intellectual exchange, transmitted through a *convivência* of language and script that within a century would disappear from the Iberian peninsula.