Words and Dictionaries. A Festschrift for Professor Stanisław Stachowski on the Occasion of His 85th Birthday, ed. E. Mańczak-Wohlfeld, B. Podolak, Kraków 2015: 297–317
doi:10.4467/K9315.25/16.15.4034

Kamil Stachowski (Kraków)
kamil.stachowski@gmail.com

Phonetic adaptation of Arabic loanwords in Argenti's Ottoman Turkish (1533). Part 1. Consonants and semivowels

0. Rationale

Arabic loanwords in Ottoman and modern Turkish are conspicuous, ubiquitous, collected and discussed in ALOT, and – perhaps for these reasons – taken for granted. But the topic is not closed; various detailed problems, such as several regarding the phonetic adaptation, remain largely unsolved and do not appear to be currently under intense scrutiny.

Some works have been devoted to this issue, e.g. Altun (2012), Korkmaz (2007), Öztekten (2013), or Yavaş (1978), but they tend to focus on modern Turkish for which, I believe, it is too soon. At any point in time, the vocabulary of a language is composed of layers inherited from different periods when different rules were in force; combined with dialectal influences, this creates a mixture that may well prove next to impossible to decompose. By starting with the earliest possible attestations, on the other hand, we give ourselves a much better chance of understanding the state that they ended up blending into.

Some research has already been made in this direction, e.g. by Stachowski M. (2012a and b).

The present paper continues this stream by presenting and preliminarily analysing the phonetic adaptation of Arabic loanwords contained in a 1533 manuscript by Filippo Argenti (ed. Rocchi 2007). Section 1 contains an introduction and a slightly more detailed look at the original spelling; section 2 the actual presentation and bits of analysis, and section 3 a brief, intermediate summary. This part is limited to consonats and semivowels; the second part will basically repeat the same scheme for vowels, and will also contain a full list of borrowings.

1. Introduction

Filippo Argenti was a Florentine diplomat in Constantinople, and the author of one of the largest Ottoman Turkish transcription texts in the Latin alphabet. The vocabulary contained in his work, more than four thousand items, was examined and published as a dictionary by L. Rocchi (2007).

Most conveniently, Rocchi included etymology in his edition. I must state clearly, however, that the etyma he gives are not strictly for Argenti's forms, but for other Ottoman or modern Turkish (literary or dialectal) words which he believes to be somehow related. I take the liberty of positing them as the sources also for Argenti's attestations, and should any of them prove to be incorrect, the fault lies with me alone.

The dictionary contains 653 units suspected of having entered Ottoman Turkish directly from Arabic. I discarded 78 when one of the following occurred: the form in the manuscript is unclear (asterisked by Rocchi 2007: 13); more than one phonetic shape of the etymon is possible; more than one reading is possible (this does not include the difficulties discussed in the excursus below, but it does include words which are only attested in an oblique form from which the nominative cannot be unequivocally reconstructed, such as *karbi* ≪ *ġarīb* (← **karip*? **karp*?), as opposed to *sultanun* < *sulṭān*, where **sulta* would be a highly improbable rendering). I did not discard words with multiple spellings if Rocchi reduced them all to one reading. Thus, this paper is based on 575 words.

Due to space limitations, and also because the focus is meant to be on phonetics and I do not want to obfuscate it by accumulating unnecessary information, I cite words without meanings. The reader is referred to Rocchi (2007). For similar reasons, I do not always adduce all of the related examples and

counterexamples, and only limit myself to what I believe are the truly relevant ones. The reader is referred to the second part of this paper, where a full list will be given.

In general, I follow Rocchi's transcription: Ottoman and modern Turkish is given in modern Turkish orthography, Arabic mostly in DIN 31635. In addition, ‹à› denotes alif maqṣūra^{h}, and ‹^{h}› denotes tā' marbūṭa^{h}. I also follow Rocchi in using ‹ä› for sounds that are spelt by Argenti either as ‹æ›, or as ‹a^{e}›, or as both ‹a› and ‹e› in different places. I accept his reading of Argenti almost automatically, though not without a little test.

Excursus: Argenti's spelling

The orthography used by Argenti is not perfect. Rocchi provides a conversion table between the original notation and his reading (2007: 5f), but specific rules are not listed, and it is evident that the correspondence is much more complicated than one to one.

Where Argenti was not very clear, Rocchi appears to have based his reading on related Ottoman and modern Turkish forms. I could not think of a direct way to confirm the accuracy of his results. The elements of the puzzle are: **1**. the Arabic source (extrapolated by me, not by Rocchi; see above); **2**. the Ottoman Turkish rendering of 1; **3**. Argenti's attempt at recording 2; **4**. forms related to 2; **5**. Rocchi's interpretation of 3 which, in some cases, was probably informed by 4. The true unknown is 2, but we cannot hope for an answer more definite than one obtained through an informed guess – such as 5. Let us try a more circumstantial approach and perform two little tests.

One particularly ambiguous sequence in Argenti's orthography is ‹ch›. It creates the impression of being used somewhat randomly, sometimes corresponding to an Arabic *h*-type consonant and Ottoman *h*, sometimes to a *ġ* and *k*, sometimes otherwise, sometimes being purely graphical. Nonetheless, a more careful examination of the correspondences makes it clear that the inconsistency is not Rocchi's. The numbers are in tab. 1, and my conclusion from them at the end of the excursus.

A related point is that Argenti's orthography makes frequent use of di-, tri- and even more -graphs. At first, collisions seem unavoidable but in fact, the only sequence that appears to require special attention is ‹sch› together with its variants ‹scch› and ‹schi›, which can be read either as *sh*/*sk* (‹s› + ‹ch›), or as *šh*/*šk* (‹sc› + ‹h›). Altogether, it occurs in seven words of Arabic origin (tab. 2).

Arabic		*ġ*	*h*	*ḥ*	*ḫ*	*k*	*kk*	*q*	*qq*
Rocchi's reading	*h*	2	17	57	34	–	–	2	–
	k	11	–	1	–	31	–	65	1

Tab. 1. Rocchi's (2007) reading of Argenti's ‹ch› in Arabic loanwords. For clarity, the table does not include the several odd instances where ‹ch› corresponds to other Arabic sounds (e.g. *ṭāliʿ* > *dalıh* ‹dalech› and ‹dalych›, *salāma*^{h} > *salamat* ‹salamacht›, or *šikāya*^{h} > *sikâat* ‹sicchiachatt›), or cases where it appears inside ‹sch› or a similar sequence (see below). The exceptional readings (*ġ* > *h*, *ḥ* > *k*, &c.) are all supported by related Ottoman and modern Turkish renderings, though in the case of *maskara* there might be some uncertainty (see below).

Arabic	Argenti	Rocchi	Related forms
ʿišq	‹eisch›	*eyisk*	Ott. *aşk olsun*, Tksh. *aşkolsun*
ʾIskandariyya^{h}	‹schienderia›	*skenderiya*	Tksh. *İskenderiye*
ishāl	‹jschal›	*ishal*	Tksh. *ishal*
maḏkūr	‹meschiur›	*mezkûr*	Ott. *mezkûr*
masḫara	‹maschara›	*maskara*	Tksh. *maskara*
muškīl	‹muschiul›, ‹muscchul›	*müskül*	Tksh. *müşkül*
taḏkara^{h}	‹teschiere›	*tezkere*	Ott. *tezkere*

Tab. 2 Arabic loanwords in Rocchi (2007) which, in Argenti's spelling, contain the ‹sch›, ‹scch› or ‹schi› sequence.

In ‹eyisch› *eyisk* and ‹muschiul› *müskül*, judging from both the Arabic etymons and the Ottoman and modern Turkish forms, the ‹sch› was more likely meant to represent *šk* than *sk*. In particular, *müskül* occurs in two spellings, with ‹sch› and ‹scch›. Rocchi considers the latter to be incorrect (puts an asterisk after it, see 2007: 13), but he does not specify the reason. Perhaps it is because ‹scch› is quite unusual in Argenti's orthography? If that were the case, one might wonder if it was not just Argenti's attempt at rendering the notation less ambiguous. Possibly, the strange spelling of *keşke*, ‹chiéschichie› stems from a similar motivation. At any rate, in the cases of ‹eyisch› and ‹muschiul›, I postulate the readings *eyişk* and *müşkül*,

respectively, and this is how they will be counted in the rest of this paper – though see also 2.1.21 below.

Also the word ‹maschara› is not entirely certain, though in this case it is not an *š* that I suspect, but a possibility of *h*. Rocchi adduces modern Tksh. *maskara*, and I can add to it the spellings ‹masqara› and ‹مسقره› in another Italian dictionary from a century later (1650; Rocchi 2011). On the other hand, in Meninski's dictionary (1680), the word is given as *mæſcharet*, *vul. maſchara*, and ‹ch› is clearly meant to represent a fricative (p. 6 of *Prooemium...*). This better matches the renderings attested in Argenti's manuscript where *ḫ* yields 38 times *h*, once Ø, and once *k* – in *maskara*.

One conclusion to be drawn from these two little tests is that Rocchi's reading (number 5), itself partially informed by the related forms (number 4), appears to be corresponding quite well to the Arabic source (number 1), and significantly better than Argenti's notation (number 3) – at least so far as *k* and *h* are concerned. Ex post facto, this can be extended to other consonants, as it will be shown in section 2 below.

Essentially, the problem has been thus reduced to the decision whether it is more likely that Argenti's notation was faithful, and he really did only encounter Ottomans who grossly distorted their Arabic and whose speech had since almost completely disappeared, or that his recording was no better than other 16th century attempts, and Rocchi's reading, consistent with the surviving and historical related forms, is in fact closer to what Argenti had actually heard. This paper is based on choosing the latter.

2. Adaptation

This section is divided into two subsections. The first one (2.1.1–2.1.28) presents the adaptations of specific Arabic sounds; the second one (2.2.1–2.2.6) particular sequences and processes.

2.1. Sounds

2.1.1. ʾ ء (hamza)

	Ø	*y*	*v*
ʾ	75	3[?]	1[?]

The three examples for *y* are *bayat* < *bā'it*, *halayık* < *ḫalā'iq*, and *layık* < *lā'iq*; the one for *v* is *suval* < *su'āl*. However, it is not at all clear whether they are actual

renderings of *ʾ*, or merely later additions, on the Ottoman Turkish ground, to avoid hiatus; see 2.2.3 below.

Besides them, there are 13 words with an intervocalic *ʾ*: two surrounded by *a*'s and *ā*'s (*barat* < *barāʾa*^{h}, *inşalla* < *ʾin šāʾa allah*), and eleven between an *ā* and an *i* (e.g. *daim* < *dāʾim*, *kail* < *qāʾil*). In the first group, *ʾ* is always dropped. In the second, with the exception of *fayda* < *fāʾida*^{h} with *tayfa* < *ṭāʾifa*^{h}, and *reha* < *rāʾiha*^{h}, all have either the *ai* sequence, or what Rocchi transcribes as *a(y)i* (*aca(y)ip* < *ʿağāʾib*, *kara(y)ip* < *ġarāʾib*, and *müla(y)im* < *mulāʾim*). This difference between the two may be illusory; see 2.2.3 below.

2.1.2. ʿ ع

	Ø	*h*	*v*	*y*	*e*	*a*	*i*
ʿ	72	2	2[?]	2[?]	2[?]	1[?]	1[?]

The two exceptions in *h* are *samah* < *samāʿ* amd *dalıh* < *ṭāliʿ*. There are five more words with *-āʿ* and also one with *-īʿ*, and they all have *ʿ* just ignored (e.g. *cima* < *ğimāʿ*, *safa* < *ṣafāʿ*, and *muti* < *mutīʿ*).

The two cases where *ʿ* is rendered as *v* are *davun* < *ṭāʿūn* and *dova* < *duʿāʾ*. There are altogether eight or nine other examples with *VʿV*, in particular four or five with *u* or *ū* in the etymon: *maona* < *maʿūn* (but cf. Ott. *mavona* and *mavuna*), *muaf* < *muʿāf*, *mamala* < *muʿāmala*^{h}, *rebi ülevvel* < *rabīʿu-l-awwal*, and the unclear one, *mavanet* < *muʿāwana*^{h} (the *v* can also be a rendering of the *w*, especially in the light of Ott. *muavenet*). The remaining four cases with *VʿV* show Ø three times and once *y* (below). One might assume that the rise of *v* is not as much a rendering of *ʿ* as it is an Ottoman intervention to avoid hiatus, but it seems that hiatus did not in fact greatly trouble Ottomans; see 2.2.3 below.

The two cases with *y* are *meymar* < *miʿmār*, and *zayıf* < *ḍaʿīf*. The first one might have more to do with how the *i* in the *iʿ* sequence is perceived than with *ʿ* per se. In the second one, *y* is most probably purely phonetic to avoid hiatus.

The two exceptional renderings as *e* are in *eyişk* < *ʿišq* and *bedeat* < *biḍʿa*^{h}; the one rendering as *a* is in *ayvaz* < *ʿiwaḍ*. As for *bedeat*, it is probably to be blamed on the auditive perception of the *dʿa* sequence – although in the other three words with the *CʿV* sequence, the *ʿ* is simply dropped (*cuma* < *al-ğumʿā*^{h}, *maşala* < *mašʿala*^{h}, and *roka* < *ruqʿa*^{h}). Similarly in the other two, the *e* and the *a* result probably from the colouring caused by *ʿ* rather than reflect the *ʿ* itself. See the five more examples with the apparently confusing anlaut of *ʿi-*: *ayar* < *ʿiyār*, *esa* < *ʿĪsā*, *ezzet* < *ʿizza*^{h}, *inaet* < *ʿināya*^{h}, and *üllet* < *ʿilla*^{h}.

The one word in *i* is *şeri* < *šarʿ*. Here, the odd rendering cannot be blamed on colouring or hiatus avoidance. However, Argenti gives the word in several forms, of which two are particularly interesting, ‹scerisíder› and ‹sceriíder›. This is reminiscent of Tksh. *cami* (< *ǧāmiʿ*) where ʿ is, in a way, preserved but only becomes visible when followed by a suffix that begins with a vowel. Apparently, *şeri* acts in a very similar way – suggesting that the *i* does not reflect the original ʿ. Rocchi adduces the meaning of *šarʿ* as 'legge religiosa', and mentions such related Ottoman forms as *şerʿ*, *şeri*, and *şerʾi*. Cf. Ott. *şariat* (vulg. *şeraʿat*) and Tksh. *şeriat* < Ar. *šarīʿa*^{h} 'legge; legge religiosa islamica'. It is tempting to declare the *i* a result of contamination. However, the latter word is given by Argenti as ‹sceraatt› and ‹sciaraat› and read by Rocchi as *şäraat*, rendering this possibility considerably less plausible and keeping the *i* unexplained.

2.1.3. *b* ب

	b	*p*	*Ø*	*m*	*u*	*f*
b	78	44	4	3	2	1

Rendering as *p* occurs mainly when the *b* is in auslaut (36 examples). This includes cases when it is a part of a consonant cluster (such as *haps* < *ḥabs*, *rup* < *rubʿ*, or *şap* < *šabb*) and even when it is only attested by Argenti in a phrase in which the next word begins with a vowel (as in *dalap ederum* < *ṭalab*). This group partially overlaps with the eight cases in which the devoicing is due to a neighbouring voiceless consonant (e.g. *mahpup* < *maḥbūb*, *spat* < *iṭbāt*, or *suphanalla* < *subḥāna'llāhi*). In two examples, an initial *b* was devoiced (*pahal* < *bāḫil* and *pelit* < *ballūṭ*), and once an intervocalic geminate (*kapan* < *qabban*). The last three cases do not seem to have a phonetic explanation, as initial *bā-* is preserved in all the remaining three cases (plus fifteen in *ba-*; e.g. *baluk* < *bāliġ*, *buhur* < *bahār*), intervocalic geminate in all the remaining five (e.g. *kubbes sehera* < *qubba*^{h} *aṣ-ṣaḫra*, *muhebbet* < *maḥabba*^{h}).

In four cases the *b* was dropped: *kapan* < *qabbān*, *şap* < *šabb*, *tabak* < *dabbāġ*, and *kürum* < *kurunb*. For the first three, see above (*şap* is clear; *kapan* and *tabak* are not). The case of *kürum* is more complicated. Although a sonorant + stop sequence is perfectly valid in Ottoman and modern Turkish (see e.g. the word *Türk* itself), the a tergo dictionary lists no more than six, and visibly much younger, words in *-mp* (Vietze, Zenker, Warnke 1975: 152: *ditiramp*, *kamp*, *katakomp*, *kramp*, *motopomp*) – followed by *zatülcenp* < Ar. *ḏātu-l-canb* (Nişanyan). It seems as if, for some reason, the Ottomans were disinclined towards the *-mp* sequence, and ready to either drop the *-p*, or to retain the phonetically inconvenient *n* in order to avoid it.

The three cases of *m* are *magmun* < *maġbūn*, *makara* < *bakra*, and *mühtan* < *buhtān*. The possibility of influence of other nasal consonants should probably be excluded; firstly because *bakra* does not have any, and secondly because there are as many as 28 words with a nasal consonant and a *b* which was, regardless, rendered as either *b* or *p* (e.g. *mektup* < *maktūb* or *sabun* < *ṣabūn*).

The two words for *u* are *daul* < *ṭabl* and *nauz* < *nabiḍ*, and the one for *f* is *sedef* < *sadāb*. It seems as if the *b* were interpreted in them as a *w*, and adapted accordingly (cf. *daur* < *tawr*, *tasfir* < *taṣwīr*), but how such a confusion could have possibly occurred, is a puzzle. See 2.2.6 below on the relation between *w* and *b*, and also 2.2.3 on the appearance of hiatus where there was not one.

2.1.4. د *d*

	d	*t*	*z*	Ø
d	50	27	2	1

The most frequent reason for rendering as *t* is that the *d* is in auslaut (19 examples). This includes cases where *d* is inside a consonant cluster, even when the other consonant is voiced (e.g. *cilt* < *ğild*, *mert* < *mard*). This group partially overlaps with the five cases in which devoicing is due to a neighbouring voiceless consonant (e.g. *kuts* < *quds*, *miktar* < *miqdār*). In five more cases, the devoicing occurred for no apparent phonetic reason, three times in anlaut (*tabak* < *dabbāġ*, *tahıl* < *daḫl*, *tellak* < *dallāk* – against eight examples with the *d* preserved, e.g. *delu* < *dalw*, *dövlet* < *dawla*^{h}) and twice surrounded by voiced sounds (*cetivel* < *ğadwal* and *halta* < *qilāda*^{h}).

The two examples for *z* are *cumazi elaher* < *ğumādà l-āḫira*^{h} and *hızmet* < *ḫidma*^{h}. Phonetically, there is no explanation for this strange rendering. Maybe a case of graphical borrowing with د being mistaken for ذ (see below)?

The one word with a dropped *d* is unusual. It is *had* < *ḥadd*, and its expected shape would have been **hat*, *ddı*. I can only offer a semi-parallel in *redd* < *radd* where, however, the Ottoman word is only attested by Argenti in the phrase *redd ederum*, i.e. in a significantly different phonetic surrounding.

2.1.5. ض *ḍ*

	z	*t*
ḍ	14	1

The one exception is *muhtur* < *muḥḍir*, which is the only case where *ḍ* stands directly adjacent to a voiceless consonant. Nonetheless, the case is not clear because a devoiced *z* would have been **s*, not *t*. A similar irregularity, though one easier to explain, can be found in *masat* < *mišḥaḍ* (see 2.1.6 below).

2.1.6. ḏ ذ

	z	*t*
ḏ	9	1

The one exception in *t* is *masat* < *mišḥaḏ*, and not at all clear. It is the only case with *ḏ* appearing in the auslaut but, firstly, voiced *z* should be a valid auslaut in Ottoman (cf. the rendering of *ḏ* below and six words in *-z* or *-ẓ* which yielded *-z*); and secondly, the devoiced version of *z* is **s*, not *t*. Nonetheless, the other dental fricative, *ṯ*, is always rendered as *t*. This opens the possibility that the final devoicing happened not after adaptation to the Ottoman system (*ḏ* → *d* or *z*), but before it: *ḏ* was perceived as *ṯ* and adapted, consequently, as *t*. See also *muhtur* < *muḥḍir* in *ḍ* above.

2.1.7. f ف

	f	*d*	*p*	*v*	Ø
f	67	1	1	1	1

The exceptions in *d* (*dalan* < *fulān*), *p* (*zurnapa* < *zarāfa*^{h}) and *v* (*şavaat* < *šafāʿat*) are all quite inexplicable. Initial *f* has been left unchanged in as many as 14 examples (e.g. *fıstık* < *fustuq*, *fursat* < *furṣa*^{h}), and intervocalic one in as many as 23 (e.g. *zarafet* < *ẓarāfa*^{h}, *zurafa* < *ẓurafā'* and *safa* < *ṣafāʿ*, *sefer* < *ṣafār*) plus one where it was in a geminate (*şaffaf* < *šaffāf*).

The one example with a dropped *f* is much clearer, as it is *kef* < *kaff* where the *f* was not only in a geminate in auslaut, but also the Ottoman word is only attested by Argenti in the phrase *kef gelurum*, i.e. directly before a consonant.

2.1.8. ǧ ج

	c	*ç*	*s*	*z*
ǧ	45	4	1	1

The devoicing happens when *ǧ* is in auslaut (e.g. *haraç* < *ḥarāǧ*, *veç* < *waǧh*).

The one exception in *s* is *mehres* < *maḥraǧ*. There is no obvious reason for this irregularity.

The only case of *z* is *zenzebil* < *zanǧabīl*. Here, a guess can be ventured that the odd *z* is due to assimilation that, perhaps, was made easier by the similarity of the resulting shape to the Turkic model of partial interfixed reduplications (as in Tksh. *sersefil* or *yemyeşil*), see also *zerzele* in 2.1.14 below.

2.1.9. ġ غ

	k	*g*	*h*	*t*	Ø
ġ	11	6	2	1	1

The rendering as *k* happens when *ġ* is in auslaut (three cases, e.g. *istifrak* < *'istifrāġ*) and, in eight out of twelve cases, when it is in inlaut (e.g. *kafil* < *ġāfil*, *kayrat* < *ġayra*^{h}).

The sound is rendered as *g* when it is in a cluster with a voiced consonant (three cases, e.g. *magmun* < *maġbūn*), in anlaut (two of twelve cases, *ganimetlik* < *ġanīma*^{h}, *garaz* < *ġaraḍ*) or, in one of two cases, between vowels (*lagum* < *laġam*).

Both renderings as *h* are in anlaut, *haşa* < *ġāšiya*^{h} and *hayr* < *ġayr*. There does not seem to be any particular phonetic cause for this deviation from *k*.

The one word in *t*, *zampt* < *şamġ*, is not at all clear; see 2.2.1 below.

The one example of dropping *ġ* is in *feraat* < *faraġa*. There is only one more example of intervocalic *ġ*, *lagum* < *laġam*. Unless a case of misspelling in Argenti's manuscript, it might be that the unusual adaptation of *feraat* is connected to the history of Ottoman/Turkish native *g* > *γ* > ‹ğ›. See also 2.2.3 below on the hiatus.

2.1.10. *h* ه

	h	Ø	*t*
h	25	12	1

The rendering as no sound occurs when the *h* is in auslaut (five of seven cases; the odd two are *tembih* < *tanbīh* and *ebleh* < *ʾablah*), part of the word *ʾAllāh* or its derivative (*alla ta ala* < *ʾAllāhu taʿālà*, *ruhulla* < *rūḥu ʾllāhi*, *suphanalla* < *subḥāna ʾllāhi*), and in four cases where there is no obvious reason for departure from the standard adaptation (*ceyiz* < *ğihāz*, *mašat* < *mašhad*, *şe(y)it* < *šahīd*, and *zar* < *zahr* – cf. e.g. *cahil* < *ğāhil*, *ishal* < *ishāl*, *šahit* < *šahīd*,[1] and *bazahar* < *bāzahr*).

The one example for *t* is *tasma* < *hašma*^{h}, which Rocchi believes to be a metathesis of Ott. *heşmet*. I am not certain that this explanation can be considered conclusive.

2.1.11. *ḥ* ح

	h	Ø	*k*
ḥ	66	8	1

The rendering as no sound seems to be caused by some as yet undetermined factors, and not by phonetics alone. It occurs in four out of eight cases when the *ḥ* is final (*carra* < *ğarrāḥ*, *saba* < *ṣabāḥ*, *sahi* < *ṣaḥīḥ*, *tespi* < *tasbīḥ*

1 Rocchi gives the meaning of *şahit* as 'testimone' and relates the form to modern Tksh. *şahit* 'teste, testimone' from Ar. *šahīd* with the same meaning. On the other hand, *şe(y)it* is translated as 'saluo' (*quando uno muore per la fede muore saluo: sceitt uluris* 'noi moiamo saluj') and compared to Balkanic Tksh. *şeyit* = literary Tksh. *şehit* 'caduto, martire', 'Muslim who died for Islam', from Ar. *šahīd* 'teste; martire; caduto per la fede'. Therefore, *şe(y)it* and *şahit* are just two adaptations of a single word. This shows that the reason for the different treatment of *h* does not lie in the phonetics of the etymon, but probably in the time or place of borrowing.

vs. *kadah* < *qadaḥ* and similar, even *sulh* < *ṣulḥ*), twice between a consonant and a vowel (*masat* < *mišḥaḏ* and *musaf* < *muṣḥaf*; but note that these are the only two cases where the consonant is a fricative; it yielded *h* when the consonant was an *l* (twice), *b*, *m*, or *r* (all once)), once intervocalically (*musaabet* < *muṣāḥaba*^{h} vs. e.g. *mastlahat* < *maṣlaḥa*^{h} or *muhebbet* < *maḥabba*^{h} and eleven other examples), and once in inlaut (below).

Out of 36 cases when *ḥ* is initial and before an *i*, in one word it was adapted to no sound (*esäp* < *ḥisāb*), in one to *h* (*hekmet* < *ḥikmā*^{h}), and in one to *k* (*kına* < *ḥinnā'*). In all the other 33 examples it is followed by *a*, *ā*, *u*, or *ū*, and rendered as *h*.

2.1.12. *ḫ* خ

	h	Ø	*k*
ḫ	38	1	1

The no sound exception is *endek* < *ḫandaq*. There are 19 cases where the initial *ḫa-* yielded *ha-*, e.g. *hamur* < *ḫamīr* or *hardal* < *ḫardal*.

As for the *k*, the odd case is *maskara* < *masḫara*. It might be that it is not actually an exception, and the correct reading is *mashara*; see section 1 above.

2.1.13. *k* ك

	k	*g*	*n*	Ø
k	53	1	1	1

The single exception in *g* is *gibrit* < *kibrīt*. There are 18 more examples for initial *k-*, and all of them have it rendered as *k*, including three where it is followed by *i* (*kiri* < *kirā'*, *kispet* < *kiswa*^{h}, *kitap* < *kitāb*), one where it is followed by *ī* (*kimiya* < *kīmiyā'*), and nine where the next consonant was also voiced (eight are sonorants, one is *kâbap* < *kabāb*).

The remaining two exceptions are both where the *k* was in a geminate, in *hakek* < *ḥakkāk* and *mehenk* < *miḥakk*. Altogether, this sequence appeared in four words, and it was preserved unchanged in the remaining two (*dükkan* < *dukkān* and *mürekkep* < *murakkab*).

2.1.14. *l* ل

	l	Ø	*r*
l	168	2	1

The two cases with no sound are both when the *l* was originally geminated, in *mähäl* < *maḥall* and *pelit* < *ballūt*. As for the former, this is the only example with a geminated *l* in auslaut. As for the latter, there are a dozen examples of intervocalic *ll* which was preserved (though four of them contain the word *'Allāh* or its derivative), e.g. *tekellif* < *takalluf* or *müsellim* < *musallam*, *tellak* < *dallāk*.

The one example with *r* is *zerzele* < *zalzala*^{h}, and probably to be explained by dissimilation (cf. *leylek* < *laqlaq*), even though in the remaining nine words that contain both *r* and *l* no such phenomenon occurred. It was maybe not without significance that the resulting *zerzele*, unlike the other nine examples, fits quite well with the native Turkic model of partial interfixed reduplications; cf. *zenzebil* in 2.1.8 above.

2.1.15. *m* م

	m	Ø
m	187	2

The only two deviations are *hamal* < *ḥammāl* and *hamam* < *ḥammām*. There are four more words with a geminated *m*, and they all have it preserved – including one case where it is in auslaut; cf. *rammal* < *rammāl*, and *hümmet* < *himma*^{h}, *meremmet* < *maramma*^{h}, *mühümm* < *muhimm*.

2.1.16. *n* ن

	n	*m*	Ø
n	108	4	1

The four exceptions in *m* are all before *b*, and the only four in which the *nb* sequence appears in the Arabic etymon: *kürum* < *kurunb* (see 2.1.3 above), *sümbüle* < *sunbula*^{h}, *tembih* < *tanbīh*, and *zambak* < *zanbaq*.

The one case of no sound is in *kına* < *ḥinnāʾ*. There are four more words with a geminated *n*, and all have it preserved, including in auslaut in one case: *cehennem* < *ǧahannam*, *cennet* < *ǧanna*^{h}, *menn* < *mann*, *unnap* < *ʿunnāb*.

2.1.17. *q* ق

	k	*g*	*h*	Ø	*y*
q	86	4	3	3	1

The four exceptions in *g* are *agbet* < *ʿāqiba*^{h}, *lagap* < *laqab*, *hegbe* < *ḥaqība*^{h}, and *rızg* < *rizq*. Phonetically exact counterexamples, where *q* would have yielded *k*, are difficult to find, but cf. *akıl* < *ʿaql*, *fakir* < *faqīr*, *takya* < *taqiyya*^{h}, *sakat* < *saqaṭ*, and *eyisk* < *ʿišq*.

The three words with *h* are *halta* < *qilāda*^{h}, *hasıl* < *qaṣīl*, and *mahrama* < (dial.) *maqrame*. Here phonetically similar counterexamples are a little easier to find; cf. *kantar* < *qintār*, *kırba* < *qirba*^{h} (plus four other words where *qi-* > *kV-*), *kasaba* < *qaṣaba*, *kabil* < *qābil* (plus 13 other words where *qa-* > *ka-*), and *akrap* < *ʿaqrab*, *mekrif* < *muqrif*.

As for the three renderings as no sound, all are where the *q* was in a geminate: *bakam* < *baqqam*, *hak* < *ḥaqq*, and *saka* < *saqqāʾ*. Two more words have this sequence, and in both it yielded *kk*: *nakkaş* < *naqqāš* and *hokka* < *ḥuqqa*^{h}.

Finally, the exception in *y* is probably a case of secondary dissimilation: *leylek* < *laqlaq*; cf. *zerzele* < *zalzala*^{h}.

2.1.18. *r* ر

	r	Ø	*l*
r	198	4	2

The cases of dropping an *r* are all from a geminated *rr*: *keret* < *karra*^{h}, *saraç* < *sarrāǧ*, *sır* < *sirr*, and *şer* < *šarr*. The last two are clear because they are final. As for the other two, cf. *carra* < *ǧarrāḥ*, *mukarrer* < *muqarrar*, *sarraf* < *ṣarrāf*, and *zerre* < *darrah* (plus three more words with a preserved intervocalic *rr*).

The two words with *l* are *kâfil* < *kāfir* and *melehem* < *marham*. The reason behind this deviation is probably not phonetics alone; see for the former *müsafir* < *musāfir* and *vafir* < *wafīr*, and for the latter *merhaba* < *marḥab*^{an}, and *mahrum* < *maḥrūm*, *rahmat* < *raḥma*^{h}.

2.1.19. *s* س

	s	*z*
s	59	1

The sole exception is *eztrafil* < *isrāfīl*. It is the only Arabic loanword in Argenti's manuscript that contains the *sr* sequence, and this fact does not considerably help explain the deviation. One might guess that *s* was first voiced between voiced sounds, and only then did the epenthetic *t* arise, as it sometimes does near sonorants (see 2.2.1). But why was it a *t* rather than a **d*? Or, alternately, why was *z* not reverted back to *s* when it found itself directly before a voiceless consonant? Argenti's notation is ‹extrafil› suggesting *ztr* quite clearly (‹x› appears 39 times in the words considered here, always for *z*); other Ottoman shapes were *esrafil* and *israfil*.

2.1.20. *ṣ* ص

	s	Ø	*z*
ṣ	51	3	1

Two out of the three exceptions where *ṣ* was dropped are the only two where it was in a geminate in auslaut, *has* < *ḫaṣṣ* and *makas* < *maqāṣṣ*. The third one is *kasap* < *qaṣṣāb*, for which the only counterpart with a geminated *ṣ* is *muhassal* < *muḥaṣṣal*.

The only case with *z* is *zampt* < *ṣamġ*, which stands against 14 examples for initial *ṣa-* being consistently rendered as either *sa-* or *se-*, e.g. *sabur* < *ṣabr* or *savap* < *ṣawāb*. There are seven more words with initial *ṣ* followed by some other vowel, all yielding *s-*, e.g. *sufur* < *ṣifr* or *sulh* < *ṣulḥ*.

2.1.21. *š* ش

	š	*s*
š	35	6

Note that I go against Rocchi about ‹eisch› and ‹muschiul› ~ ‹muscchul›, and read them with *š* instead of *s*. See section 1 above.

The six odd cases are *istah* < *'ištihā'*, *masat* < *mišḥaḏ*, *müsteri* < *muštari*^{n}, *sikâat* < *šikāya*^{h}, *sükürla-* < *šukr*, and *tasma* < *hašma*^{h}. Exact phonetic counterexamples are not attested in Argenti's manuscript; the closest ones are *aşık* < *ʿāšiq*, *eyisk* < *ʿišq* (see above), *haşa* < *ġāšiya*^{h}, *haş haş* < *ḫašḫāš*, *maşat* < *mašhad*, *şarat* < *'išāra*^{h}, and *şöhret* < *šuhra*^{h}. It seems that the unusual adaptation cannot be linked to any of the following sequences: *aš*, *iš*, *ši*, *šu*. For *uš*, the only other example is the somewhat problematic *müşkül*. But the sequences *šḥ*, *šm*, and *št* do not appear in any more words, and therefore they could be blamed for causing this aberration. Yet, the specific phonetic reason for this eludes me for now, and if it were found in the future, it would still leave *sikâat* and *sükürla-* unexplained.

2.1.22. *t* ت

	t	*d*	*n*	Ø
t	65	2	1	1

The two words in *d* are *dan* < *taʿn* and *daur* < *tawr*. They are countered by 28 cases of initial *t* yielding *t*, including 25 where it is followed by *a*; e.g. *tayin* < *taʿyīn*, *tazir* < *taʿzīr*, *tebrit* < *tawrā*^{h}, and *terk* < *tark*. However, see also *davun* < *ṭāʿūn* and *daul* < *ṭabl* below.

The one case in *n* is *intifak* < *'ittifāq*, and the one with no sound is *keten* < *kattān*. These are the only two Arabic words with a geminated *t*. As for the former, see also 2.2.1.

See also 2.2.5 below on tā' marbūṭa^{h}.

2.1.23. *ṭ* ط

	t	*d*	*k*
ṭ	23	4	1

The four exceptions in *d* are *dalap* < *ṭalab*, *dalıh* < *ṭāliʿ*, *davun* < *ṭāʿūn*, and *daul* < *ṭabl*. There are also four words in *ṭa-* which yielded *ta-* (*tabak* < *ṭabaq*, *takya* < *ṭaqiyya*^{h}, *Tarabulus* < *Ṭarābulus aš-šāmi*, *taraf* < *ṭaraf*) and two in initial *ṭ* followed by another vowel (*tayfa* < *ṭā'ifa*^{h}, *tufan* < *ṭūfān*). See also the exceptional *dan* < *taʿn* and *daur* < *tawr* above.

The one odd word in *k* is *aktar* < *ʿaṭṭār*. The only other example for a geminated *ṭ* is *battal* < *baṭṭal*.

2.1.24. *ṯ* ث

	t
ṯ	5

The five words are *abes* < *ʿabaṯ*, *espap* < *'aṯwāb*, *mesel* < *maṯal*, *sevr* < *ṯawr*, and *spat* < *iṯbāt*. See also *ḏ* above.

2.1.25. *w* و

	v	Ø	*b*	*f*	*p*	*u*
w	46	6	2	2	2	2

The six words where *w* was dropped are *af* < *ʿafw*, *coap* < *ǧawāb*, *cöher* < *ǧawhar*, *müzevir* < *muzawwir*, *nöbet* < *nawba*^{h}, and *täfaut* < *tafāwut*. Only the first of them is clear; there is one more word with *-Cw* and it, too, has this sequence modified (*delu* < *dalw*). As for the others, cf. *savap* < *ṣawāb*, *cevlan* < *ǧawlān*, *evvel* < *ʾawwal* (and three more words with a geminated *w*), *dövlet* < *dawla*^{h}, though for *tafāwut* there does not appear to exist a convincingly similar counterexample. In two cases a hiatus was created which was not there in the original; see 2.2.3 below.

The two cases with *b* are *hebala* < *ḥawāla*^{h} and *tebrit* < *tawrā*^{h}. Cf. *hava* < *hawā'* (plus four examples for *awā*), *sevr* < *ṭawr*. I cannot think of any particular phonetic reason for fortition here; see also the examples for *p* below.

The two with *f* are *fetfa* < *fatwà* and *tasfir* < *taṣwīr*. These are two of five cases where *w* was preceded by a voiceless consonant; the remaining three all have unusual renderings: Ø above and *p* below.

The two cases with *p* are *espap* < *ʾaṭwāb* and *kispet* < *kiswa*^{h}. Apart from *fetfa* and *tasfir*, perhaps the closest to them are *cetivel* < *ǧadwal* and *tezvir* < *tazwīr*. The voicelessness is clear, but fortition is much less so, just as with the two examples in *b* above. See also *daul* and *nauz* in 2.1.3 above.

Lastly, the two with *u* are the already mentioned *delu* < *dalw* and *daur* < *tawr*. Cf. also *kavs* < *qaws* and *cevz* < *ǧawz*. See 2.2.3 on hiatus.

See also 2.2.6 below on the relation between *w* and *b*.

2.1.26. *y* ی

	y	Ø
y	41	12

The dozen cases of dropped *y* are not quite clear.

Four of them are when the etymon ended in the *iyya*^{h} sequence (e.g. *vesiyat* < *waṣiyya*^{h}). There is only one more example of a geminated *y*; it does not end in this sequence, and it has the *yy* preserved: *mukayyat* < *muqayyad*.

The remaining eight cases are peculiar. Three had the *āya* sequence (*inaet* < *ʿināya*^{h}, *sikâat* < *šikāya*^{h}, and *vilaet* < *wilāya*^{h}), and there are no more examples for it in Argenti's manuscript. Further, there is one in *ayb* (*cep* < *ǧayb*), one in *ayh* (*alehi* < *ʿalayhi*), one in *ayk* (*salamalek* < *salām ʿalayk*), one in *ayz* (*hazaran* < *ḫayzurān*), one in *iya* (*haşa* < *ġāšiya*^{h}), and save the last one, for neither of those sequences are there more examples. The words with *iya* are *cariye* < *ǧāriya*^{h} and *terbiyet* < *tarbiya*^{h}. Note the three words where an unnecessary hiatus was created; more on this in 2.2.3 below.

2.1.27. *z* ز

	z	*d*	*s*	Ø
z	46	1	1	1

The one case in *d* is *mahaden* < *maḫzan*. There are no other examples of the *ḫz* sequence. However, the cluster is split in the Ottoman form, and there are as many as eleven examples for intervocalic *z* rendered as *z* (in one case, it is no longer intervocalic: *zar* < *'izār*; some of the others are *bazahar* < *bāzahr*, *mizan* < *mīzān*, and *mülazimet* < *mulāzama*^{h}).

The sole exception in *s* is *sokak* < *zuqāq*. This is the only case with initial *zu-*, but there are 17 words in *za-*, *zā-*, and *zi-*, and they all have the *z* preserved.

The only case with a dropped *z* is *veledizina* < *waladu 'z-zinā'*, and it is one of two words which had a geminated *z* in the etymon. The other one is *ezzet* < *'izza*^{h}, with no change made to the cluster. There are in fact three more Ottoman words in Argenti's manuscript, with a geminated *z*, even if stemming from *ḍḍ* or *ẓẓ*: *lezzet* < *laḍḍa*^{h}, *hazzet-* < *ḥaẓẓ*, and *mutavazza* < *mutawaẓẓā'*.

2.1.28. *ẓ* ظ

	z	*d*	*t*
ẓ	10	2	2

The two words with *d* are *hadır* < *ḥāẓir* and *nadar* < *naẓar*. To the former, the closest example is probably *intizar* < *'intiẓār*, but to the latter it is certainly *nazar* < *naẓar*. It is the same Arabic word, which simply has two reflexes in Ottoman Turkish. This clearly shows that phonetics is not the sole factor that determines the final rendering but sadly, it does not show what those other factor or factors are.

The two renderings as *t* (in *lahta* < *laḥẓa*^{h} and *matlama* < *maẓlima*^{h}) are the only examples where the *ẓ* is adjacent to a consonant. (From the Arabic perspective, *vaz* < *waʿẓ* is another one, but Ottomans tended to simply entirely ignore the ʿ (see above).)

2.2. Other

2.2.1. Epenthesis

Epenthesis is attested in a number of Argenti's examples, but here we will limit ourselves to consonants. There are 66 cases of final *-t* where there was *tā' marbūṭa*^{h} in Arabic (so much as this can be considered epenthesis; see 2.2.5 below), about seven of *v* and *y* inserted to neutralize hiatus (2.2.3 below), five less definite examples (below), and finally one entirely unclear (*eyişk*, see 2.1.2 above).

The five words with extra consonants are *eztrafil* < *isrāfīl* (see 2.1.19 above), *mastlahat* < *maṣlaḥa*^{h}, *testlim* < *taslīm*, *zampt* < *ṣamġ*, and *zurnapa* < *zarāfa*^{h}.

The first three contain the *sl*/*sr* sequence. There are three more words with it, and only one has it preserved (*asıl* < *ʾaṣl*, *ıslah* < *iṣlāḥ* and *Mısır* < *Miṣr*). There are further three words with the *zl*/*ẓl* sequence (preserved in one, *nuzla* < *nazla*^{h} as opposed to *kazıl* < *ġazl* and *matlama* < *maẓlima*), and one word with a *zr* sequence that was not there in the Arabic original (*hazret* < *ḥaḍra*^{h}).

As for *zampt*, there are no more examples for *mġ*, final or otherwise. Perhaps the closest phonetically is *mḥ* in *lamḥa*^{h} which resulted in a metathesis: *hamla*. There are no examples for *mh*, *mḫ*, *mr* or the same with an *n*.

Finally, *zurnapa* is entirely unclear.

Thus, the only visible tendency is for the *sl*/*sr* sequence to be removed – through epenthesis of *t* or, if in auslaut, of a vowel.

2.2.2. Geminates

In total, there are 85 cases of geminated consonants in Arabic etymons; 54 are also rendered as geminates in Ottoman.

The most common case when a geminated consonant is shortened, is when it is in auslaut. This happens in ten out of twelve examples. The two exceptions are *menn* < *mann* (Argenti only gives *menn ederum* but there exists an independent Ott. *menn*[2]), and *muhimm* < *mühümm*. The last two are Argenti's only examples with a geminated *m* or *n* in auslaut. Possibly, two more words should be counted here, *hazzet-* (< *ḥaẓẓ*) and *redd et-* / *redd olun-* (< *radd*), which however only appear coupled with a verb in Argenti's manuscript, and therefore with the geminate not in auslaut.

The remaining 21 words are less clear. They are listed above sub *b* (2 examples), *k* (1), *l* (1), *m* (2), *n* (1), *q* (2), *r* (2), *ṣ* (1), *t* (2), *ṭ* (1), *w* (1), *y* (4), and *z* (1) above. All are intervocalic (with the half-exception of *veledizina* < *waladu ʾz-zinā'*), and if we exclude the cases in auslaut, then for each of these sounds but *y* and *t*, there are as many or more examples where the geminated consonant was preserved. It is surprising that of all sounds, it is *tt* that was shortened in both cases, seeing how often the sequence appears in native words if, admittedly, on morpheme boundaries only (*dörtte*, *gitti* &c.). Nonetheless, there do not seem to be enough similarities to speak of a common theme.

---

2 See e.g. Zenker (1866–76: 881; after Rocchi 2011), and probably also à Mesgnien Meninski (1680) who attests *menn* '*i.a.* debilitare, diminuere; benevolum, benignum propitium eſſe; benefacere; beneficentia, donum, gratia, favor', while the meanings given in Rocchi are 'absoluo, perdono, quieto, fo fine et quitanza, lo libero del debito'.

2.2.3. Hiatus

Hiatus only appears in two Arabic etymons, *ʿaib* (> *ayıp*) and *ġaib* (> *kayıp*) – or three, if we count *ʾin šāʾa allah* since it is rendered as one word in Ottoman (*inşalla*). In all of them, it has been removed.

It has also been removed in further thirteen words where a *ʾ* or *ʿ* (or, in one case, an *h*) was dropped. It is not always clear what method was actually used in each specific case (cf. especially the examples in 2.1.1 and 2.1.2 above), but they certainly include epenthesis (seven cases, e.g. *davun* < *ṭāʿūn*, *layık* < *lāʾiq*), contraction (four cases, e.g. *barat* < *barāʿa*^{h}, *inşalla* < *ʾin šāʾa allah*), and shift to a semivowel (two cases, *fayda* < *fāʾida*^{h} and *tayfa* < *ṭāʾifa*^{h}).

In four words, it is not certain whether an anti-hiatus consonant was or was not inserted. They are ‹aggiajpp› aca(y)ip < *ʿağāʾib*, ‹charaipp› *kara(y)ip* < *ġarāʾib*, ‹mulaim› *müla(y)im* < *mulāʾim*, and ‹scieitt› *şe(y)it* < *šahīd*.

In two more cases a repair, of sorts, was attempted: *alla ta ala* < *ʾAllāhu toʿālà*, and *rebi ülevvel* < *rabīʿu-l-awwal*.

Overall, the number of removed hiatuses and potential hiatuses, is between fifteen and twenty-one.

It might come as a surprise, then, that hiatus is present in as many as twenty or twenty four Ottoman words where there was none in Arabic (the unclear four being the ones with a bracketed *y*). The sequences are *aa* (five cases, e.g. *musaabet* < *muṣāḥaba*^{h}, *şavaat* < *šafāʿat*), *ae* (two cases, *inaet* < *ʿināya*^{h} and *vilaet* < *wilāya*^{h}), *ai* (five or nine cases,[3] e.g. *daim* < *dāʾim*, *kail* < *qāʾil*), *au* (four cases, e.g. *daul* < *ṭabl*, *täfaut* < *tafāwut*), and one example for each of *ao* (*maona* < *maʿūn*), *ea* (*bedeat* < *biḍaʿa*^{h}), *oa* (*coap* < *ğawāb*), and *ua* (*muaf* < *muʿāf*). Most of these hiatuses result from ignoring the Arabic *ʾ* or *ʿ* (eleven words); three cases are due to the dropping of *y*; further four are due to *w* and *b* (two each), and lastly two are due to *ġ* and *ḥ* (one each); see sub respective letters for all four.

2.2.4. Metathesis

There are three cases of metathesis: *hamla* < *lamḥa*^{h}, *nalet* < *laʿna*, and *tercibe* < *tağriba*^{h}. It is difficult to say what may have caused this unusual adaptation. The only word that is phonetically similar to any of these three is *mane* < *maʿnà* (and *nanä* < *naʿna*ʿ, but this one is obviously of no use here). The *lam* sequence appears in three more words (always yielding *lVm*), but *mḥ* and *ğr* or similar clusters occur nowhere else. In other surroundings,

3 The unclear four are *aca(y)ip* < *ʿağāʾib*, *kara(y)ip* < *ġarāʾib*, *müla(y)im* < *mulāʾim*, and *şe(y)it* < *šahīd*.

l and *r* are quite abundant (see 2.1.14 and 2.1.18 above) but they are always kept in their original places.

2.2.5. Tāʾ marbūṭa^h

	V			*Vt*				Ø
	a	*ä*	*e*	*at*	*ät*	*et*	*it*	
a^h	35	1	15	23	1	39	–	1
ā^h	1	–	–	–	–	1	2	–

The odd case of no sound is *cumazi elaher* < *ǧumādà l-āḫira^h*. There is no obvious phonetic reason for this inconsistency; cf. especially *ahırat* < *al-āḫira^h*.

The two exceptions in *-it* are *divit* < *dawā^h* and *tebrit* < *tawrā^h*. Words in *-ā^h* are very rare in general; among those in *-a^h*, endings *-ra^h* and *-wa^h* occur thirteen and four times, respectively, and yield consistently *a* and *e*, e.g. *halvat* < *ḫalwa^h*, *kayrat* < *ġayra^h*.

As for a more general picture, rendering as a front or back vowel, and rendering with or without *-t* are approximately equally frequent (resp., 57:59 and 52:66 cases). Also, neither of these two correlates with the Arabic short or long *-a* in auslaut (resp., $p = 0.3598$ and 0.6294; Fisher's exact test). Interestingly, however, rendering with or without *-t* appears to be correlated with harmony ($p = 0.0001908$, and if the two cases in *-it* are excluded, 0.0003487; FET). This result puzzles me.

2.2.6. *w* vs *b*

A special relation between *w* and *b* may seem like a rather banal observation to make, but the case is slightly more complicated in Argenti's material.

The usual rendering of *w* is *v* or *f*, sporadically Ø or *u*. However, it yielded *b* or *p* in four out of sixty words (*espap* < *ʾaṭwāb*, *hebala* < *ḥawāla^h*, *kispet* < *kiswa^h*, and *tebrit* < *tawrā^h*). Phonetically, there does not seem to be anything particularly special about them, cf. e.g. *hava* < *hawā'*, *tasfir* < *taṣwīr*, *tezvir* < *tazwīr*, or *sevr* < *ṭawr*. Perhaps these borrowings are older than the rest, or maybe they penetrated into Argenti's manuscript from the speech of a person from outside Istanbul?

The available material does not allow one to make definite claims, but the connection may in fact run deeper. Against 129 cases of *b* yielding the understandable *b*, *p*, *m* or an occasional Ø, there are three where it was turned into *u* or *f*: *daul* < *ṭabl*, *nauz* < *nabiḍ*, and *sedef* < *sadāb*. There are not very many phonetically similar counterexamples but cf. *ebleh* < *ʾablah*, *kıbla* < *qibla^h*, *kabil* < *qābil*, *tabi* < *tābi*, or *savap* < *ṣawāb*, *harap* < *ḫarāb*. The first two can perhaps be explained by a hypothetical intermediate stage with a [v] or [β] *([daβuł], [naβuz]), but for *sedef* I can think of no excuse.

3. Summary

Argenti's material is rich but it is not complete, its interpretation is not always certain, and it is not in itself sufficient for definite conclusions to be drawn. However, three of the points raised in this paper are in a way interesting, and I want to repeat them here for convenience, together with a more general observation.

Firstly, hiatus has proven to be surprisingly not problematic for 16th century Ottomans. Some of the details remain unclear, but attempts to remove it are visible in just as many or even fewer words than it was actually created in on the Turkish ground, where there was none in Arabic. (Admittedly, it was often only ʾ or ʿ that separated the vowels, but this could have been quite effortlessly fixed with a *v* or *y*.) See 2.2.3.

Another unexpected result is the correlation between the adaptation of tāʾ marbūṭaʰ and vowel harmony. When tāʾ marbūṭaʰ is rendered as a vowel, harmony tends to be back, and when it yields *-Vt* vowels tends to be front – much more frequently than would be expected if the two features were in no way, directly or indirectly, connected. See 2.2.5.

The third point is just a single word, *feraat* < *faraġa*, which may prove helpful in dating the native Turkish change *g* > *γ* > ‹ğ›. See 2.1.9.

Overall, the adaptation of Arabic consonants in Ottoman as recorded by Argenti can be characterized as follows. With the sole exception of the anyway rare *ṯ*, all consonants are rendered in more than one way. The most diversified are the adaptations of ʿ, but then the interpretation is not always straighforward in this case. Almost equally diversified, and less ambiguous, are *b* and *w*. On the opposite end of the scale are *ḍ*, *ḏ*, *m*, *s*, *š*, *ṭ*, and *y*. Almost all of the Arabic phonemes have one, very clearly dominating Ottoman counterpart.[4] However, a phonetic surrounding that could be shown to dictate the specific adaptation appears in the vast majority of cases to be impossible to establish. Counterexamples with a very similar or identical surrounding can be found for most words with unusual adaptations, and prove not infrequently to be actually more numerous. The only aberrations that can be clearly linked to phonetics, are the simplest and the most predictable ones: devoicing in auslaut and before voiceless consonants, and shortening of final

4 This can be expressed more precisely using normalized Shannon entropy – which, here, should be interpreted as a measure of uncertainty about the Ottoman Turkish rendering that is associated with specific Arabic sounds. It has the range from 0 to 1, and for more than three quarters of the sounds discussed here, it is below 0.5.

geminated consonants; yet even in these areas exceptions occur. One might hypothesize about these odd cases being older borrowings or dialectal shapes, but the answers must wait for more material from the period to be extracted, and its phonetics to be analysed.

References and abbreviations

ALOT = Stachowski S. 1975–1986.
Ar. = Arabic.
Nişanyan = Nişanyan S., *Nişanyan Sözlük. Çağdaş Türkçenin Etimolojisi*, www.nisanyansozluk.com.
Ott. = Ottoman.
Tksh. = Turkish.

Altun H.O. 2012. Türkiye Türkçesindeki Arapça Alıntı Kelimelerde Ünlü Uyumsuzlukları. – *Türkiyat Araştırmaları Dergisi* 32: 57–71.

Korkmaz F. 2007. Alıntı Kelimelerin Türkçeleşme Sürecinde Bazı Arapça Kelimelerin Görev Değişikliğine Uğraması Üzerine. – *İlmî Araştırmalar* 23: 103–112.

à Mesgnien Meninski F. 1680. *Thesaurus Linguarum Orientalium Turcicæ, Arabicæ, Persicæ* […]. Vienna.

Öztekten Ö. 2013. Osmanlı Türkçesindeki Arapça Alıntılarda Yapımlık Ses Değişiklikleri. İlletli Kökler. – Şirin User H., Gül B. (ed.), *Yalım Kaya Bitigi. Osman Fikri Sertkaya Armağanı*. Ankara: 1–12.

Rocchi L. 2007. *Ricerche sulla lingua osmanlı del XVI secolo. Il corpus lessicale turco del manoscritto fiorentino di Filippo Argenti (1533)* (= *Near and Middle East Monographs. New Series* 5). Wiesbaden.

Rocchi L. 2011. *Il Dizionario Turco-Ottomano di Arcangelo Carradori (1650)*. Trieste.

Stachowski M. 2012a. On the consonant adaptation of Arabic (and some Persian) loan-words in an Ottoman Turkish dictionary by Arcangelo Carradori (1650). – *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 129 *supplementum*: 373–381.

Stachowski M. 2012b. Some features of the vowel adaptation of Arabic loan-words (along with a few remarks on their consonants) in an Ottoman Turkish dictionary by Arcangelo Carradori (1650). – *Folia Orientalia* 49: 491–502.

Stachowski S. 1975–1986. *Studien über die arabischen Lehnwörter im Osmanisch-Türkischen* (= *Prace Komisji Orientalistycznej PAN* 13, 15, 17, 19). Wrocław – Warszawa – Kraków – Gdańsk – Łódź 1975, 1977, 1981, 1986.

Vietze H.-P., Zenker L., Warnke I. 1975. *Rückläufiges Wörterbuch der türkischen Sprache*. Leipzig.

Yavaş M. 1978. Borrowing and its implications for Turkish phonology. – *Kansas working papers in linguistics* 3: 34–44.

Zenker J.Th. 1866–1876. *Türkisch-arabisch-persisches Handwörterbuch*. Leipzig.