MAREK STACHOWSKI Jagiellonian University, Cracow stachowski.marek@gmail.com

# ON THE CONSONANT ADAPTATION OF ARABIC (AND SOME PERSIAN) LOAN-WORDS IN AN OTTOMAN TURKISH DICTIONARY BY ARCANGELO CARRADORI (1650)

**Keywords:** Ottoman Turkish, Turkic, Arabic, Persian, transcription texts, Transkriptionstexte, Arcangelo Carradori, linguistic contact, lexicography, historical phonetics

#### **Abstract**

In this study an analysis of the phonetic adaptation of Arabic and Persian loan-words in Ottoman Turkish is continued (for the vocalic part of the analysis see Stachowski M. [forthcoming]). Five phenomena are presented in the context of the general Turkic phonetic evolution. These are: [a] palatalization of  $(-)kE->(-)\check{c}E-$ ; [b] varying an analytic m->b- and b->m-; [c] despirantization of f>p; [d] epenthetic n; [e] shortening of geminates.

## 1. General remarks

Luciano Rocchi's (2011) publication of a 17<sup>th</sup> century Ottoman Turkish dictionary by Arcangelo Carradori made it possible to observe certain features in the phonetic adaptation of Arabic loan-words in (mostly, as it seems, colloquial) Ottoman Turkish. The importance of his attestations and the problems connected with their analysis were discussed briefly in a previous article that focused on vowel adaptation (Stachowski M. forthcoming). There is thus no need to repeat these arguments.

In this article five phonetic phenomena are presented. The starting point is generally a set of Arabic (and sometimes Persian) words attested in Carradori's dictionary. However, the processes observed more often than not also concern native Turkic words, so that a general Turkological background is discussed as well.

# 2. Consonant changes

### 2a. Palatalization of Ar., Pers. (-)kE- > Tksh. (-)čE-1

No Arabic example of this change could be found in Carradori's dictionary. However, in another source from the 17<sup>th</sup> century there is the following example:

Ott. *çursi* (or perhaps better: *çürsi* ? – M.S.) ‹csurszi› 'podium, teacher's desk' < Ar.  $kurs\bar{\imath}$  'id.; chair; throne' > modern Tksh.  $k\bar{u}rsi\sim k\bar{u}rs\bar{u}$  'podium; chair, department' (N. Illésházy's 1668 dictionary, cited after ALOT II 54).

Oddly enough, Persian loan-words displaying this change are, unlike Arabic ones, recorded in Carradori's work, e.g.:

Ott.  $\check{c}a\check{g}at$  «ciaghat» ~  $k\hat{a}\check{g}id$  «chiaghed» ~  $k\hat{a}\check{g}at$  «chiagat» ~  $k\hat{a}hat$  «chiahat» 'sheet of paper, card' (196) < Pers.  $k\check{a}\check{g}a\underline{d}$  ~  $k\check{a}\check{g}i\underline{d}$  'paper' > modern Tksh.  $k\hat{a}\check{g}id$  'paper'.

Ott.  $cemence < ciemencieh > kemence < chiemence > 'violin with three strings' (104)^2 < Pers. <math>content = content =$ 

Ott. hunçarlık (or maybe:  $h\ddot{u}nçarlık$ ? – M.S.) (hunciarlich) (but (ḥnkārlk) in Arabic script) 'belonging to a sovereign' (179) < \* $h\ddot{u}nk\hat{a}rllk$  < \* $h\ddot{u}nk\hat{a}r$  'sovereign' > modern Tksh.  $h\ddot{u}nk\hat{a}r(l\imath k)$  id.

Only one indigenous Turkic word exemplifying this change could be found:

Ott. kivircik «chiuirgich» ~  $\varsigma ivircik$  (or possibly:  $\varsigma \"overcik$  ~  $\varsigma \"ovircik$ ? – M.S.) «ciouergich» 'curly, frizzy' (216).

As there is no reason to assume that the Ottoman Turkish pronunciation of *k*-in Persian loan-words differed significantly from that of the Arabic *k*- the lack of Arabic examples of this type in Carradori's dictionary could be nothing more than a coincidence.

However, another aspect of this phenomenon seems to be even more intriguing. The kE > cE palatalization is a usual assimilation device of Ottoman Turkish loanwords in the Balkan Slavonic languages, e.g. Serbian cebap < Tksh. cebap < Kebab (Stachowski S. 1973: 116), Serbian cebap < Tksh. cebap < Tksh. cebap < Tksh. cebap < (op. cit. 193), Serbian cebap < Tksh. ce

The symbol E stands for any palatal or palatalizing vowel, e.g. e in  $\langle ke \rangle = [ke]$ ,  $\hat{a}$  in  $\langle k\hat{a} \rangle = [ka(:)]$ .

Double- and three-digit numbers in parentheses always refer to the pages in Rocchi (2011). Notations in \(\cdots\) are either Carradori's Latin-script renderings of Ottoman Turkish words or Latin-script transliterations of a traditional Arabic spelling.

like *çursi*, *çemençe*, and so on, this view now appears to be far less certain than was previously thought.

The problem is whether or not we are actually capable of differentiating between words palatalized in Turkish dialects and borrowed with their dialectal  $\varsigma$ - into Serbian on the one hand, and those that are only palatalized in Serbian, on the other. At any rate, the Turkish-speaking population must have, it would seem, been aware of the k-  $\sim \varsigma$ - alternation because hypercorrect k- forms are also known, as is e.g. Ott.  $\varsigma$ -ember  $\langle$ -cember  $\rangle$  and its more recent hypercorrect variant  $\rangle$ -kember  $\rangle$ -band (Rocchi 2009: 69)  $\langle$ -Pers.  $\check{g}$ -ambar id.

### 2b. Varying anlaut nasality: Ar. m- > Ott. b-; Ar. b- > Ott. m-

Only one example for each of these changes is attested in Carradori's dictionary:

Ott. *bember* 'bember' 'pulpit (in the mosque)' (78) = *member* id. (G. Molino's 1641 dictionary) < Ar. *minbar* [-mb-] id. > modern Tksh. *minber* [-mb-] id.

Ott. *makra* (maqra) 'pulley, block' (237) < Ar. *bakra* id. ~ Syrian Ar. *bakara* id. (Tietze 1958: 265sq.) > Ott. *makara* (macchara) id. (F. Argenti's manuscript of 1533; cited after Rocchi 2007) = modern Tksh. *makara* id.

In the case of the Ott. *bember* one might hypothesise that we are dealing with the change m - mb > b - mb which can be, at least theoretically, interpreted as both assimilation (m - b > b - b) and dissimilation (m - m > b - m). The reason for the change was without doubt the rather typical Turkish tendency to avoid sonorants in the word-initial position.

The situation with Ott. mak(a)ra is entirely different. There is no m or b in the non-initial syllable, and the change Ar. b- > Ott. m- runs against the principle of avoiding word-initial sonorants. Thus, the Ott. mak(a)ra seems to have adjusted perfectly to a whole group of Turkic words with the m-  $\sim b$ - alternation in the syllable-initial position. Various examples of such changes can be observed throughout the Turkic linguistic world, e.g. in Chulym  $simicak \sim sibicak$  'two-year old horse',  $simicak \sim sibicak$  'little finger' (Pomorska 2001: 98); Chulym  $simicak \sim sibicak$  'twe-year old horse',  $simicak \sim sibicak \sim sibicak$  'little finger' (Pomorska 2001: 98); Chulym  $simicak \sim sibicak \sim sibicak \sim sibicak \sim sibicak \sim sibicak of sibicak of$ 

Various explanations for this phenomenon were proposed over the course of the history of Turkic linguistics, beginning with a suggestion that this consonant was originally a type of affricate  ${}^*m^b$  (or maybe  ${}^*m^b$ ), rather than just  ${}^*m$  or  ${}^*b$  (Ramstedt 1957: 74sq.), which is the only explanation that allows for word pairs like Ott.  $buz \sim Yak$ . mus 'ice'. However, neither this conjecture nor the idea of assimilation ( ${}^*b - N > m - N$ )³ or dissimilation ( ${}^*m - N > b - N$ ) can convincingly explain why some Arabic, Persian and Russian loan-words have also undergone this change, even if they show no nasal consonant in a non-initial syllable and cannot possibly reflect Proto-Altaic shapes with an affricate  ${}^*m^b$  or  ${}^*m^b$ , e.g. Arabic bakra (see above); Ott.  $bala \sim {}^*m^b$ 

The symbol N stands for any nasal consonant.

mala 'headscarf' < \*bala < Pers. vāla 'a sort of silk cloth' (Tietze, Tekin 1994: 161); Tuvinian batrās ~ matrās 'mattress' < Russ. matrac id. (Schönig 2002: 263), Tobolsk-Tatar müts 'oven' < Russ. peč' id. (Anikin 2003: 448).<sup>4</sup>

## 2c. Despirantization of the Ar. f > Ott. p

In view of the fact that the Turkic languages did not originally possess a labial spirant f, as well as that this sound is, in loan-words, more often than not replaced in Anatolian Turkish dialects by a stop p, it is quite remarkable that only one such example is attested in Carradori's dictionary:

Ott.  $p\ddot{a}na$  (panah) 'abuse, misuse' (152) < Ar.  $fan\bar{a}$ ' id. > modern Tksh. fena 'bad'.

On the other hand, the consonant -*p*- in the Turkish reflexes of Ar. *maṭbaḥ* 'kitchen' still remained occlusive in 17<sup>th</sup> century Turkish, as confirmed by Ott. *motpak* 〈motpach〉 id. (246) ~ *mutbak* (G. Molino's 1641 dictionary) [in reality certainly pronounced -*tp*-] = *mutpah* or *mutpak* 〈mutpach〉 id. (F. Argenti's data of 1533) (all records cited after Rocchi 2011: 246) even if it is a spirant in modern Tksh. *mutfak* id.

The situation could be interpreted to the effect that Ottoman Turks had learned to pronounce the spirant f by the mid-17<sup>th</sup> century but despite this were still treating it as a foreign consonant that sometimes could be replaced by p (as, e.g. in  $p\ddot{a}na$ ) but generally did not occur in lieu of p in loan-words (cf. motpak).

The latter formulation is possibly a touch too rigorous. The example of Ott.  $sulf\ et$ -(1sg.aorist: (sulf ederum)) 'to (be)calm, soothe' (300) < Ar. sulh 'peace' (> standard Ott. sulh id.) shows that the consonant f could, at least as the second member of an intervocalic consonant cluster, replace another consonant in an Arabic etymon. Such cases, however, are extremely rare, and any attempt at generalization appears premature.

Ott. *ness* (ness) 'greed, lust, craving, desire' (256) < Ar. *nafs* id. does not belong to this group. Instead, it results from the simple assimilation of the Ar. *-fs* > Ott. *-ss*, i.e., *-s#* and -s + C- but -ss + V, cf. [2e].

#### 2d. Epenthetic n

The n epenthesis is an irregular phenomenon that only rarely affects Arabic loanwords. The most frequent examples in Carradori's dictionary are native derivatives with the nomina agentis suffix -(i)ci, as e.g. Ott. yudanci (iudangi)  $\sim yudaci$  (iudagi)  $\sim yutuci$  (ywtwğy), etc. 'swallower, swallowing' (357) < yut- 'to swallow', amongst others.

<sup>&</sup>lt;sup>4</sup> The word pair Mo. *bolğa*- 'to depend' ~ Khakass *molğa*- id. probably displays the same relationship, i.e. Mo. *b*- > Khakass *m*- (Rassadin 1980: 39); however, one might consider a Pro-Altaistic solution like: Proto-Alt. \**m*<sup>b</sup> ≫ Mo. *b*- ~ Khakass *m*-, as well.

<sup>&</sup>lt;sup>5</sup> The symbol  $\ddot{a}$  is used by L. Rocchi (2011) to render a vowel that is generally expected to be e but is spelled (a) in Carradori's dictionary.

Interestingly enough, one does not find the form *yudancı* in a historical dictionary of Ottoman Turkish -(*i*)*cı* derivatives which, instead, adduces a 1499 record *yudunıcı* ⟨yud-un-iği⟩ id. (Stachowski S. 1996: 179). As -(*i*)*cı* derivatives are at the same time "amazingly frequently subject to elision in Carradori's dictionary" (e.g. Ott. *unutcı* ⟨vnutgi⟩ 'oblivious' (332) < \**unutcı* < *unut*- 'to forget', see Stachowski M. forthcoming: § 2e) their composition and phonotactic evolution deserve further detailed investigation.

Only one Arabic example in this group is perfectly clear: Ott. *manzul* (manzul) 'dismissed' (237) < Ar. *ma'zūl* id. > modern Tksh. *mazul* id.

Another word, namely Ott. *musarif* seems at first sight to display no -*n*- at all. However, its evolution, as based on Rocchi's (2011: 250) suggestion, can be summarized in the following manner:

Ott. musarif (musarif) ~ munsarif (mnṣrf) 'author' (250) < \*munsanif < \*munsannif < Ar. muṣannif id. > standard Ott. musannif id.

The form \*munsannif will, unlike Ott. manzul, have resulted from the anticipation of a nasal, i.e., Ar. -C - nC - NC - nC which is in essence just a case of assimilation.

In contrast to this word, Ott. *daimanlı* ‹daimanlı · (~ *daimalı* · (daimalı) 'endless' and Ott. *daimanlık* · (daimanlich) (~ *daimalık* · (daimalich) 'continuation' (114) seem to display a simple -n- epenthesis in derivatives of Ott. *daima* 'always'. However, this simple explanation does not in truth hold water since Ott. *daimalı(k)* and *daimanlı(k)* reflect two different starting points: Pers. *dāyimā* 'always' and its source: Ar. *dā'iman* id. (Tietze 2002: 548).

Yet another case is Ott. *getavrez* (ghetaures) 'centaur' (161)<sup>6</sup> = *gentavres* (ghentaures) id. (1641 dictionary by G. Molino) = *kentavriş* (1687 *Onomasticon* by F. Meninski) id. (all attestations adduced after Rocchi 2011: 161) < Greek *kéntauros* id. One might suspect a hypercorrect elision of -*n*- if Turkish-speaking informants were more or less aware of the existence of the -*n*- epenthesis. This does not, however, appear very realistic because the epenthesis seems to have been a relatively rare phenomenon. In addition, the word *centaur* will certainly have extremely seldom occurred in Turkish conversations. In this situation, Luciano Rocchi's conjecture that the form without -*n*- is probably nothing other than a careless notation appears to be the best solution. However, the -*n*- is lacking in both the Latin and the Arabic spelling – can both be a matter of carelessness?

## 2e. Shortening of geminates

<sup>&</sup>lt;sup>6</sup> Rocchi (2011: 161) identifies the Latin notation with the Arabic one in that he writes: "ghetaures = kṛāwrz" which seems to be two spellings of the same phonetic shape *getavrez*. I would prefer to read the Arabic-script notation with *k*- in Turkish. Nevertheless, both forms lack an -*n*-.

afv (i.e., = [af:]) ('fw) 'forgiveness' (50sq.) that has, unlike *ness*, partially retained the geminate before an anlaut vowel in certain words: Ott. *afv iste*- (1sg.aorist: ⟨afu isderum⟩) 'to ask for forgiveness' (50) but *af et*- (1sg.aorist: ⟨af ederum⟩) 'to forgive' (51). This situation is the opposite to what exists today. The modern distribution of *af* and *aff* also depends on the character of the next element but in another way: [α] *af* becomes *aff* if the next element is a possessive suffix (e.g. *affim* 'my forgiveness') or an auxiliary verb (*affet*- 'to forgive', *affolun*- 'to be forgiven/condoned') but [β] it remains *af* if followed by an inflectional suffix (e.g. dative *af+a* albeit the accusative form usually is *aff+i*) or a non-auxiliary verb (e.g. *af iste*- 'to ask for forgiveness').

Is, thus, the impression correct that the distribution of geminate and degeminated consonants in a word-final position was either different or at least not yet stable in the year 1650?

Word-medial geminate consonants are even more interesting. Let us first study the examples recorded in Carradori's dictionary:

## 1) -*k*- group:

Ott. dükân (duchian) 'shop' (133) < Ar. dukkān id. > modern Tksh. dükkân id.

Ott. hoka (hhoqa(h), hocha) 'pot, cup' (177) < Ar. hukka id. > modern Tksh. hokka id.

Ott. *hokabaz* 〈hochabaz〉 *~ hokbaz* 〈hoqbaz〉 'illusionist, magician' (177sq.) < Ar. *ḥuķ-kabāz* id. > modern Tksh. *hokkabaz* id.

Ott. *moşekil* (moscechil) 'appearance, look' (246) < (? Pers. <) Ar. *muşakkil* 'wellshaped' > standard Ott. *müşekkel* id.

Ott. *mürekeb* ['ink'], in: *basma mürekeb* «basma murechieb» 'printing ink' (75) < Ar. *murakkab* 'composite' > modern Tksh. *mürekkep* 'ink'.

## 2) *-d-* group:

Ott. hidet \hidet \'anger' (177) < Ar. hidda(t) id. > modern Tksh. hiddet

### 3) -*b*- group:

Ott. *lübe* (*lube*?), in: *boğaz lübesi* (*lubesi*?) ‹boghas lubesi›<sup>8</sup> 'base of the neck' (234) < \**lübbe* < \**libbe* < Egyptian Ar. *libba* id. (l.c.) = Classical Ar. *labba* id. > standard Ott. *lebbe* id.

The Ott.  $zubon \, \langle zubon \rangle \sim zubun \, \langle zubun \rangle$  'a sort of jacket' (369) is a distinct word which admittedly displays a single -b- when compared with the Ar. -bb- in  $\check{g}ubba$  'a sort of outer garment, gown'. However, the degemination occurred in Venetian rather than in Turkish: Ott.  $zubon \sim zubun$  (= modern Tksh. zubun)  $< *zubon \sim *zubon < Venetian <math>zupon \sim zipon = standard$  It. gubbone id., demin. < gubba id. < Ar.  $\check{g}ubba$  (Stachowski S. 1996: 118).

This is a modern Turkish synchronic perspective, of course. Historically, one would say that the original *aff*  $\langle f w \rangle$  remains unchanged in  $[\alpha]$  and becomes degeminated in  $[\beta]$ .

<sup>8</sup> Rocchi (2011: 234) puts *lubbe* in the first place but the etymology points to *lübbe* as the more probable variant.

### 4) -z- group:

Ott. *lezetli* 'dezetli' 'tasty, spicy' (233) ~ *lezetlik ile* 'dezetlich ile' 'tastily, with taste' (l.c.) ~ *lezetle*- (1sg.aorist: 'dezetlerum') 'to flavour, season, add some seasoning' (l.c.) < \**lezet* < Ar. *laddat* id. > modern Tksh. *lezzet* id.

### 5) -ν- group:

Ott. evel (euuel) ~ evvel (euuel) '(shortly) before' (177) < Ar. awwal 'first' > modern Tksh. evvel 'before'.

Ott. *kovet* «chouet, couet, qouet» 'power, strength' (223) ~ *kuvet* «quuet» id. (229) < Ar.  $k\bar{u}wwat$  d. > modern Tksh. *kuvvet* id.

## 6) -l- group:

Ott. zavalık ‹zaualich› 'poverty' (366)<sup>9</sup> < \*zavallık < \*zavallılık id. < \*zavallı 'poor' < \*zaval < Ar. zawāl 'death, destruction' > Ott. zaval > modern Tksh. zavallı 'poor', zavallılık 'poverty'.

Ott. *zelilik* (selilich) 'weakness' (366) < \**zelillik* < Ar. *dalīl* 'humbled, abased; obedient' > modern Tksh. *zelil* id.

No degemination is observed in the trill rr or the nasal stop nn (no records with -mm- could be found in the source). These consonants always seem to retain these geminates, e.g.:

Ott. sarraf (sarraf) 'money-changer' (286) < Ar. sarrāf id. > standard Ott. sarraf id.

Ott. *minnet* ·minnet · 'thanks, indebtedness' (245) < Ar. *minnat* id. > modern Tksh. *minnet* id.

Even if the degemination of -*ll*- was exemplified above, the examples in group (6) differ considerably from all the other records in groups (1)-(5) in that they only show the -*ll*- > -*l*- change at morpheme borders between a borrowed Arabic word and a native Turkish suffix +*lık*. By contrast, the -*ll*- group never seems to be subject to degemination if it occurs in the word-medial position in an Arabic word, as for instance in the following case:

Ott. millet (millet) 'nation' (245) < Ar. millet id. > modern Tksh. millet id.

Thus, a Degemination Resistance Hierarchy (= DRH) can be established:

- 1) The velar stop k (numerous examples);
- 2) The spirants v, z and non-velar stops b, d (fewer examples);
- 3) The lateral *l* (degeminated only in derivatives);
- 4) The trill r and the nasal stops m (?) and n (no known examples).

Rocchi (2011: 366) also allows the reading zavallık which is not totally out of the question but, on the other hand, there is no suggestion of a necessity of reading -ll- here.

The lower the group number (= the DRH index), the more frequent the degemination. Actually, group [3] and group [4] can even be considered one group consisting of sonorant consonants. At the opposite end of the DRH axis one finds the velar voiceless stop k. This seems to suggest that the most important factor in the degemination process is the voicedness of the consonants involved. Another important factor, it would seem, is the degree of "consonantism", that is syllabic consonants have a higher DRH index. The mutual relationship between the two factors can only be understood when more examples are found and analysed in the future.

Nevertheless, even these insights can be used in a discussion of more general topics in Turkic linguistics, as well as in arguments for or against certain claims. When wondering at W. Hesche's suggestion that the Turkic postposition *syŋaru* 'towards' goes back to \**syŋar+ru* (< *syŋar* 'direction' + directive suffix +*ru*) I have argued that "the -*r r*- sequence (as in Tksh. *nar receli* 'pomegranate marmelade' [...]) is never shortened to -*r*-" (Stachowski M. 2012: 131). These concerns are now supported by the fact that both the sandhi -*r r*- sequences and the word-medial -*rr*- in Arabic loan-words in Turkish have the highest possible DRH index.

#### References

ALOT = Stachowski S. 1977. Studien über die arabischen Lehnwörter im Osmanisch-Türkischen. [Teil II: K-M]. Wrocław etc.

Anikin A.E. 2003. *Ètimologičeskij slovaŕ russkich zaimstvovanij v jazykach Sibiri*. Novosibirsk. Pomorska M. 2001. The Chulyms and their language. An attempt at a description of Chulym phonetics and nominal morphology. – *Türk Dilleri Araştırmaları* 11: 75–123.

Ramstedt G.J. 1957. Einführung in die altaische Sprachwissenschaft, I: Lautlehre. Helsinki. Rassadin V.I. 1980. Mongolo-burjatskie zaimstvovanija v sibirskich tjurkskich jazykach. Moskva.

Rocchi L. 2007. Ricerche sulla lingua osmanlı del XVI secolo. Il corpus lessicale turco del manoscritto fiorentino di Filippo Argenti (1533). Wiesbaden.

Rocchi L. 2009. Il lessico turco nell'opera di Bernardino Pianzola. Materiali per la conoscenza del turco parlato di fine Settecento. Trieste.

Rocchi L. 2011. Il dizionario turco-ottomano di Arcangelo Carradori (1650). Trieste.

Schönig C. 2002. Anmerkungen zum  $b \sim m$ -Wandel in den modernen Türksprachen. – Hauenschild I. et al. (eds.) *Scripta Ottomanica et Res Altaicae*. [FS B. Kellner-Heinkele]. Wiesbaden: 257–275.

Stachowski M. 2012. On the origins of the Turkic directive suffixes *-jar(y)* and *-sar(y)*. – *Studia Etymologica Cracoviensia* 17: 127–135.

Stachowski M. [forthcoming]. 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).

Stachowski S. 1973. Fonetyka zapożyczeń osmańsko-tureckich w języku serbsko-chorwackim. Wrocław etc.

Stachowski S. 1977.  $\rightarrow$  ALOT.

- Stachowski S. 1996. Zur Geschichte des osmanisch-türkischen *zıbın* 'eine Art Obergewand'. *Studia Etymologica Cracoviensia* 1: 117–124.
- Tietze A. 1958. Direkte arabische Entlehnungen im anatolischen Türkisch. Eckmann J. et al. (eds.) *Jean Deny Armağanı*. Ankara: 255–333.
- Tietze A. 2002. Tarihi ve etimolojik Türkiye Türkçesi lugatı. İstanbul Wien.
- Tietze A., Tekin T. 1994: Tarama Sözlüğü üzerine bazı açıklamalar, II. *Türk Dilleri Araştırmaları* 4: 159–169.