



CODEX CUMANICUS

CODEX CUMANICUS, a manuscript of eighty-two paper leaves, measuring approximately 20 x 14 cm, preserved in the Biblioteca Nazionale of the cathedral of San Marco in Venice and comprising principally vocabularies and texts of the Northwest Middle Turkic language of the Cumans, or Komans, recorded in Latin script. The Cumans (Russian Polovtsy) were a nomadic Turkish tribe who in the 12th and 13th centuries, together with the Petchenegs (Pečenegs), established a state in southern Russia, north of the Black Sea, which was destroyed by the Mongols. Those Cumans who were not driven farther west, into Hungary, were absorbed into the Tatar tribes of the 14th-century Kipčhak (Qïpčaq) state. In the codex the language is called *comanicum* (*chomanicho*) in Latin but *tatar til* (or *tatarçe*) in Turkish. The first line of the text bears the date 1303 (MCCCIII). From the many obvious errors, however, it is certain that this is not the original work but a copy made at some later date. It came into the possession of the Italian poet and bibliophile Petrarch (Francesco Petrarca), who presented it as part of his library to the Venetian Republic in 1362. There, apart from a short excursion to the Bibliothèque Nationale in Paris, under Napoleon, it has remained ever since. It contains (p. 72) a list of months in which January is equated with the Muslim month Şafar and December with Moḥarram (*mugarā*); this coincidence occurred most precisely, in the period in question, in 726/1325-26. The original was thus compiled in the first three decades of the 14th century. The first part, comprising 110 pages of a Low Latin-Persian-Turkish (Cuman) dictionary, was probably written by Italian merchants or Franciscan monks engaged in missionary work in the region between the lower Volga and the Crimea. The second part (pp. 111-164,



a few left blank or containing scribbles or later Italian text), evidently added by German-speaking monks, consists of Cuman-German and Cuman-Latin word lists, grammatical paradigms, and Cuman translations of texts intended for use in religious services. For example, the Lord's prayer begins (p. 126 l. 27): [Atamis kim kōcta sen, algiszle bulsun senig atig, kelsin senig hanlechin] *Atamiz kim kōkte-sen, alyišli bolsun seniņ atij kelsin seniņ xanlixij.*

The question why a form of Persian should appear in a work evidently compiled for practical purposes in this region has not been completely answered. Most probably Persian was used by the Turkish informants of the Italians as a lingua franca for trade in the east. As there is evidence that their Persian vocabulary had to some extent undergone a phonological change in common with their mother tongue, namely the development of medial and final *-d(-) > -y(-)* (e. g., Old Turk. *qaḍyu* “sorrow” > CC [kaygi] *qayyi* “mesticia”; cf. Persian *karīdār* “buyer” > CC [ghriaar] **karīyār* “emtor,” *pedar* “father” > CC [piar] *pīyar* “pater,” *bad* “bad” > CC [bad] but also [bay] “malus,” [sabte bay] **sakt bay* “pessimus, very bad”), it must have been acquired, at least in part, at a considerably earlier date. The use of Latin letters for the transcription of the two Asian languages is somewhat inconsistent, even within the “Italian” part of the manuscript, and thus the Persian material, comprising some 1,500 words, cannot entirely bear the weight of some of the phonetic and dialectological theories that have been imposed upon it by recent editors. The copying errors have already been mentioned: e.g., [psen] (next to [farxen]) for *[pser], i.e., *pesar, farzand* “filius, son”; [targos] for *[cargos] = *kargūš* “lepus, hare”; [aut] for *[art] = *ārd* “farina, flour.” There remains the possibility of others not yet recognized, or indeed recognizable. Many words have been corrupted beyond recognition, for example, [gap] for “humilitas, *lowness” (Cuman [youaslic] *yovašlik* “*docility”). Persian phonemes foreign to Latin (and Italian), in particular, are rendered in a variety of ways. The fricatives, for example, may be written as follows: /š/ generally as s: [squista] = *šekasta* “fractus, broken,” x: [padisa, padixa] = *pādešā(h)* “imperator, emperor” (cf. [inginoxus] for “ingeniosus”), but also z: [ziazm ab] = **čašm(a-ye) āb* “fons, spring”; /ž/ as x: [xala] = *žāla* “rosata, dew”; /č/ as ç: [ça] = *čah* “puteus, well,” ci: [ciana] = *čāna* “melonus, chin,” or z [zasiden] = *čašīdam* “gustavi, I tasted”; /j/ as i: [ania] = *ānjā* “illuc, there,” y: [yaga] = **jāygā(h)* “locus, place,” gi: [giou] = *jow* “ordeum, barley,” ci: [ciul] = *jol* “coperta, saddlecloth,” or z: [taz] = *tāj* “corona, crown,” /k/ as ch: [chatir] = *kāter* “voluntas, will,” gh: [ghana] = *kāna* “domus, house,” c: [car] = *kar* “asinus, donkey,” k: [kom] = *kām* “crudus, raw,” g: [dragt] = *derakt* “arbor, tree,” or h: [mehamden] = *mīkandam* “irrideo, I ridicule”; /ğ/ as



g: [gan] = *gam* “cura, care,” gh: [ghat pisiar] = *gāyat besyār* “nimis, very much,” or c (for *g?): [carbīl] = *garbīl* “sedacius, sieve”; and so on. The vowels are equally ambiguous: [e], for example, stands for */a/ in [ser] = *sar* “capud, head” (beside [sarmoxa] = *sar-mūza* “argorge, shoe”), */i/ (or /e/) in [asech] = *āseq* “inamoratus, in love,” */ay/ in [peda] = *paydā* “visible,” */ē/ in [nec] = *nēk* “bonus, good,” [sped] = *sapēd* “albus, white,” and frequently for /ā/ with *emāla*, e.g., [anden] = *andēm* “membrum, limb,” [lehef] = *leḥēf* “culter, quilt.” The *majhūl* vowels, however, were evidently at least in the process of change. For example, /ē/ was apparently preserved in the modal particle [me-] **mē-* (but [mi-] consistently before /ā/: [miaen] **mīāyam* “venio, I come,” etc.), the prefix [be-] *bē-* “without,” and elsewhere, as in [omad, umet] **ōmēd, omēd* “hope,” [pes] **pēs* “mixellus, leprous” and **pēš* “ante, before,” [tex] **tēz* “prestus, quick.” It had, however, become /ī/ before many sibilants, for example, [ysan] *īšān* “ille, they,” [andissa] **andīša* “cogitatio, thought,” and generally before nasals, [binmar] **bīmār* “infirmus, ill,” [nim rox] **nīmrōz* “nona, noon.” In view of the many contradictions, for example, [cho] **kōh* “mons, mountain” but [ruj] **rūy* “vultus, face,” [osta] **ōstā(d)* “doctor, teacher” but [pust] **pūst* “pellis, skin,” [rox] **rōz* “dies, day” but [çualdus] **jowāldūz* “aquus grossum, packing needle,” it is difficult to evaluate the fate of /ō/, especially as [o] frequently also represents /o/, e.g. [ghosch] = *košk* “sichus, dry.” Thus, for all its interest, there are both too little uniformity and too much uncertainty in the material to make it a reliable source for any precise phonological or lexicographical study of medieval Persian.

BIBLIOGRAPHY

Text: K. Grønbech, *Codex Cumanicus . . . in Faksimile*, Monumenta Linguarum Asiae Maioris 1, Copenhagen, 1936.

Cuman Turkish: A. von Gabain, “Die Sprache des Codex Cumanicus,” *Philologiae Turcicae Fundamenta* I, Wiesbaden, 1959, pp. 46-73 (with detailed bibliography).

K. Grønbech, *Komanisches Wörterbuch*, Copenhagen, 1942.



Persian: A. Bodrogligeti, *The Persian Vocabulary of the Codex Cumanicus*, Budapest, 1971 (review by D. N. MacKenzie, in *JRAS*, 1973/1, pp. 64-66).

D. Monchi-zadeh, *Das Persische im Codex Cumanicus*, Uppsala, 1969 (review by D. N. Mackenzie, in *Oriens* 25-26, 1976, pp. 391-96).

C. Salemann, "Zur Kritik des Codex Cumanicus," *Bulletin de l'Académie Imperiale des Sciences de St. Petersbourg*, 6th ser., 4, 1910, pp. 943-47.

(D. N. MacKenzie)