

Peter Bakker & Heinz Werner Wessler

THE ETYMOLOGY OF THE ROMANI QUOTATIVE *XÁCĒ*, *XÁTĒ* 'I/YOU/HE SAID'

0. INTRODUCTION

The Welsh Romani word *xácĕ*, *xátĕ* is used when a person is quoted literally, e. g. in:

- (1) *wantsésa mō gad? xáč'ī rānī. aua, xáč'o lovinákerō.*  
do.you.want my shirt said.the lady yes said.the butler  
'Dost thou want my smock?' asked the lady. 'Yes', replied the butler' (Sampson 1926: 176; his translation)

This word seems to be virtually limited to the so-called Northern<sup>1</sup> dialects of Romani (Bakker 1999), the language of the Roma or Gypsies, which belongs to the Indic branch of Indo-European. It is attested not only in Welsh Romani, but also in Russian Romani, Ukrainian Romani, Baltic Romani (Lev Tcherenkov, p.c.; Mānuš 1997a, 1997b) and in English Romani (Smart/Crofton 1875), in (virtually) the same form and the same function. A Kalderash dialect, of the Vlax branch, shows a similar form. In this paper we will propose an etymology for this quotative element, which we will tentatively gloss as QP, quotation particle.<sup>2</sup>

1. FORM VARIATIONS AND FUNCTIONAL DIFFERENCES OF *XÁCĒ* IN ENGLISH ROMANI, LATVIAN ROMANI, RUSSIAN ROMANI, WELSH ROMANI AND SERBIAN KALDERAŠ ROMANI

Both English Romani (ER), Latvian Romani (LR), Russian Romani (RR) and Welsh Romani (WR) use more than one form.

Smart/Crofton (1875) list two different forms for ENGLISH ROMANI, viz. *hótcher* and *hótchi*, of which the second seems to be more frequent: *hótcher mé* 'I said' and *hótchi yov* 'he said'. The material is too scarce to deduct a functional difference between the two.

LATVIAN ROMANI uses *xamé* 'relating to one's own words' and *xací* 'relating to someone else's words' (Mānuš 1997a: 62–63). Both are considered particles by Mānuš.

1 The Northern dialects comprise roughly all dialects traditionally spoken in Western Europe, from Finland and Scandinavia to Spain and Italy and from Britain to Germany and Austria. Dialects spoken in Poland, North Russia and the Baltic states are also Northern.

2 The abbreviations used in the glosses are: ABL – ablative, ART – article, DAT – dative, DEM – demonstrative, EMPH – emphatic marker, F – feminine, M – masculine, NEG – negative, negation, NOM – nominative, OBL – oblique, PART – participle, PERF – perfect, PL – plural, PRES – present, PST – past, QP – quotation particle, quotative, SG – singular.

- (2) *me phučháva látir: ›Kaj tu, xamé, džása?‹*  
 1SG.NOM ask.PRES.1SG from.her where you QP you.go  
 'I ask her where she is going' (litt.: I asked her, where – I said – are you going?)
- (3) *joj rakírla: ›me xačí davá nadžindžòm.‹*  
 she say.PRES.3SG 1SG.NOM QP this know.NEG.PST.1SG  
 'She says she didn't know this' (litt.: She says, I – she said – did not know this)

The spoken words are quoted literally, and the particles apparently appear in the middle of the quote. They can also be repeated throughout the quote in Latvian Romani (Mānuš 1997b: 341).

- (4) *me rakiráva léske: ›Me xamé nasòm xamé jáσαι xamé diliní,*  
 1.NOM. I.speak to.her 1.NOM QP I.am.not QP such.a QP crazy.F  
*sir tumé xamé dufinéna xamé mándir‹*  
 as.if you.PL QP think.2PL QP 1.ABL  
 'I say to him: ›I am not as crazy as you think of me‹'
- (5) *Jov rakírla mánge: ›mò dád xačí isí xačí kaná nasvalö*  
 he.NOM say.3 to.me my.M father QP is QP now ill  
*tè navidžála xačí nikáj xačí khēréstir avrí.‹*  
 and NEG.he.goes QP nowhere QP house.ABL out  
 'He says to me: ›My father is ill now and he is going nowhere from the house‹.'

Welsh Romani uses *xáčē* and *xâte*. The form *xâte* is used only once in Sampson's examples (1926: 175, see example 6 below). The form *xáčē* is used in past and present and in singular and plural, probably in all persons. In some cases the definite articles *o* (masculine) and *i* (feminine) and the personal pronouns *jov* 'he', *joj* 'she' and *jon* 'they' cliticize to *xáčē*, giving respectively *xáč'ō*, *xáč'ī*, *xáč'ov*, *xáč'oi*, *xáč'ón*.

- (6) *›pirā ī stiga, ŷukel, xâtē mē leskī.*  
 open the gate dog QP 1SG.NOM to.him  
 '›open the gate, dog‹, cried I to him.'
- (7) *›mīrī sī 'doiyā, xáč'ō Ĵak.*  
 mine.F is DEM.F QP.ART.M Jack  
 '›That one (girl) is mine‹, cried Jack.'

It is not clear what the difference in function (if there is one) is between *xáčē* and *xâte*.

Russian Romani (Patkanov 1911: 122, 123, 124) has *xáci*, *xáci-ta* and *xáme*. The *-ta* suffix is probably the common Romani emphatic suffix. Some examples (spelling as in source, Sampson 1926: 176, 177, quoting Patkanov) are:

- (8) ›*ná kinen*,‹ *chăči-ta*, ›*sivoné grén: bibachtalé avéna*‹.  
 NEG buy.2PL QP grey.OBL.PL horse.OBL.PL unhappy.PL they.come  
 '›Do not buy grey horses,‹ said he, ›they will bring us ill luck‹.'
- (9) ›*ná-akija*,‹ *chăči*, ›*but, mre romá*‹.  
 NEG-so QP much, my.PL Gypsy.PL  
 '›Not so much, my Gypsy friends,‹ quoth he.'
- (10) *So, cháme, za bíbacht?*  
 what QP ?? unhappy  
 '›What misfortune is this?‹, I wondered.'

The element *xame* seems to be used for first person and *xate* for third person, as in Latvia.

Only the *xăcē* and *xătē* forms are used as quotatives in WR, but WR also has the *xâme* form (Sampson 1926: 177), but it is used with a different function. In WR it is interpreted as meaning 'pretending'. In fact, there are three forms for this in WR: *xâme*, *xoi* and *xâme-xoi*, with apparently the same function.

- (11) ›*Aua*,‹ *p'endóm maia xâmē lakī*.  
 yes I.spoke 1SG.EMPH PRETEND to.her  
 '›Yes,‹ said I jestingly to her.'
- (12) *Ĵă, bă, xâmē kin ī gres*.  
 go mate PRETEND buy the horse.OBL  
 'Go, mate, pretend to buy the horse'

The meaning „pretending“ may very well be a secondary development from the meaning „I say/said“ and „I think/thought“. The main difference between this form in WR and the others is that the quotes here are not literally. Sampson (1926: 177) suggests a link between WR *xâme* and the same particle in Russian Romani.

Even though the root is otherwise only attested in Northern dialects, a cognate form is apparently in use in Kalderaš Romani as well, which belongs to the Vlax branch and which shows a lot of Rumanian influence. Boretzky (1994: 269) lists a word *xáci*, translated 'allegedly', for a variety of Kalderash spoken in Serbia. He does not give an etymology. It does not seem to be reported from other Kalderaš varieties (e. g. Calvet 1993; Demeter/Demeter 1990):

- (13) *xáci phenel kə kadja sî*.  
 allegedly he.says that so is  
 'He allegedly says it is so'

The shape of the Kalderaš word is identical and even though the meaning is slightly different, it is certainly close to both other meanings, roughly 'speaking' and 'pretending'.

In brief, the particle (Mānuš) or defective verb (Sampson) *xače* with variants is used in direct quotes, either after the first word or constituent in the quote, or preceding or following it. It is found virtually only in Northern dialects, but not in all of them.

## 2. PAST PROPOSALS FOR THE ETYMOLOGY OF XAČĚ

Despite the relative rareness of the word, the etymology of *xače* has been the subject of discussion of at least four earlier researchers.

Smart/Crofton (1875) suggest a link with a Rumanian Romani word *hiotosaro* 'to utter loud cries' and Paspatian Romani *khuyázava* 'to call', but these are implausible, because of form, meaning and the unlikelyhood of a Rumanian loanword in this dialect (see note 3).

Sampson (1926: 175–176) noted "etymology obscure", and he rejected the proposed Rumanian and Greek etymologies as he found them unconvincing. A derivation from English *quoth* is also rejected.

Gilliat-Smith (1927: 29) proposed to compare it with "Rumanian Gypsy" *-iče*. It is a form also found, according to Gilliat-Smith, in the Romani dialect of the "Varna sedentary sieve-makers".

Winstedt (1944) discussed the etymology of English Romani *hoči* 'he said'. Winstedt relates it to Rumanian *iče*, connecting it to Banat and Transylvanian Rumanian. In these Rumanian dialects the first vowel would be schwa-like, and /h/ or /x/ can be added in vowel-initial words. There are four reasons to object to its Rumanian origin. First, it is not documented for any form of Rumanian (as Winstedt himself admits). Second, there are otherwise no Rumanian loanwords in the Northern dialects or in Welsh Romani.<sup>3</sup> Third, the word cannot be found in any Vlax dialect (those spoken by Roma which have Rumanian influence). Fourth, it is the element *xa-* which recurs cross-dialectally. The element *-če* is one of four variants: *-me*, *-či*, *-če* and *-te*. In no case is there an *i-* prefix.

Neither of these etymologies is convincing. Based on the stress pattern it is most likely of Asian origin. It has final stress. The shared Romani lexicon is of Indic, Iranian, Armenian and Greek origin, and is stress-final, whereas European words have penultimate stress (see Boretzky 1989; Bakker 1997).

---

3 The Rumanian loanwords listed by Sampson (1926) are almost all general Balkan words. Nevertheless, there are a handful of words in several Northern dialects which must be from Rumanian, but these are very few.

## 3. AN INDIC ETYMOLOGY

There is no convincing etymology of these forms yet. The broad geographical distribution makes it likely that it is to be traced down to an Asian rather than a European source.

Cross-linguistically, quotative particles are often derived from verbs meaning 'to speak'. Modern Indo-Aryan languages have two basic groups of verbal stems meaning 'to speak' that go back to the OIA (i. e. Old Indo Aryan) roots *(ā-)khyā* (Turner 1966, nr. 1040–1043) and *kath-* (Turner 1966: 2702–2708). This is indeed the clue to the etymology of Romany *xăcē/xate*. We will first discuss the shape of the word.

Romani initial /x/ in Indic roots often goes back to a OIA aspirated unvoiced velar /kh/. In some cases /x/ goes back to a consonant cluster of an unvoiced velar and a sibilant: /*[sibilant]-k/* or /*k-[sibilant]/*. The common initial OIA /kʃ/ for example often developed into /kh/ in Prakrit, and afterwards to /x/ in Romani. Some examples (all from Sampson 1926) are:

- |   |   |
|---|---|
| (14) S. <i>kha-</i> 'cavity, hole',               | WR <i>xev</i> 'hole';   |
| S. <i>khan-</i> 'to dig',                         | WR <i>xand-</i> 'to dig (up)';                                  |
| S. <i>khura</i> 'hoof, claw',                     | WR <i>xurr</i> 'hoof'.  |
| S. <i>kṣāma</i> 'scorched', <i>kṣā</i> 'to burn', | WR <i>xāvalī</i> 'embers, cinder';                              |
| S. <i>kṣudra</i> 'small, petty',                  | WR <i>xuredō</i> 'short, small' (other dialects <i>xudro</i> ). |

These two sources for initial Romani /x/ are the most common ones. There is a third source for Romani /x/: in some cases it goes back to a OIA initial unvoiced velar /k/, which was first aspirated for reasons discussed below, and then developed into /x/ in Romani. This third possibility offers a convincing clue to the etymology of *xate*. The phenomenon of aspiration of originally non-aspirated, single OIA stops appears typically in Dardic and (West-) Pahari languages, which show special developments of the feature of aspiration. Masica (1991: 203ff.) records cases of "historically unmotivated" voiced aspirates in these languages, which are independent of cases of aspiration motivated by a following /r/ or /s/ or by a lost \*s + consonant in pre-OIA. This is a very common phenomenon in most New Indo-Aryan languages. Cases like Old Indo-Aryan *jānāti* 'knows' > Kalasha *jhon-*, Phalura *jhan-* demonstrate a peculiar aspiration of originally unaspirated initial stops in roots. A case like Old Indo Aryan *bandhati* > Lakasha *bhon-* corresponds to a devoiced variant in European Romani *phand-* (Turner 1959 [1975]: 381).

Turner (1959 [1975]: 381) analyses cases of transference of aspiration in Romani from the last consonant of a root to the first consonant, a phenomenon that had already been recognized earlier by Wackernagel in general (OIA and Greek), and by Miklosich (1878) and Sampson (1926) for Romani. Turner, however, notes mostly cases of exchange of aspiration between a voiced initial stop and a voiced non-initial aspirate. He lists a number of cases. He notes a single

case that does not fit into this pattern: Romani *phuc-* 'to ask', since the initial is unvoiced in Old Indo-Aryan (*prcchati*), and the stem retains the internal aspirate. Most of the cases demonstrated something like a modified Grassmann's law,<sup>4</sup> which had been noted by Allen (1957) already, the so called "transference of aspiration" as noted above. Masica (1991: 204) provides the following example: *khatā*, going back to Old Indo Aryan *kathā* 'story'. These cases are apparently not limited to exchanges of aspiration between voiced stops. Aspiration is transferred to voiced as well as to unvoiced initials. Masica (1991: 204) gives examples for this phenomenon not only from Pahari and Dardic, but also from Konkani, down from the Southwest coast of India. Alternatively, the aspiration shift in Old Indo-Aryan *kath* may also be a result of interference from forms going back to the Old Indo Aryan *khyā*.

Taking all this evidence of historical phonology together, it is clear that Romany /xəčə/, /xate/ correspond to the Old Indo Aryan root *kath-* 'to speak', i.e. *kathāyati* (Turner 1966, nr. 2703) 'converses'. Derivatives of this verb exist in the Modern Indic languages, e. g. Western Hindi *kahnā*, Rajasthani *kainō*, Gujarati *kahēvun*. Other forms are derived from OIA (*ā-*)*khyā-* like Lahnda *ākhan*, Panjabi *-ākḥṇā*. Forms in /kh/ and /k/ are mixed up in Modern Indic languages, derived from the two OIA roots.

In short, the root *kath-* transferred its aspiration to the initial consonant, yielding *khat*. Following the evidence we cannot decide whether this shift of aspiration has happened in the Prakrit ancestor of Romani or later, parallel to a similar development in other Modern Indic languages. Later the /kh/ changed into /x/ in Romani, as attested in several other words. The /t/ palatalized to /č/, a process natural enough not to warrant a special explanation. Thus, *xace* is an Indic root in Romani, one that has not been identified as such.

#### 4. A SIMILAR FUNCTION IN AN INDIAN LANGUAGE?

Quotative particles are very common in the languages of India. The quotative or reported speech particles are clearly an areal phenomenon (Kāchru 1979; Masica 1991: 403), used in both the Dravidian and Indic languages. Probably under the influence of Dravidian they spread to the Indic languages. These quotative markers have a variety of forms. Many of these quotative particles are derived from a verb meaning 'to say'. This is also the case in Romani, as shown above.

Actually, we also find quotatives in India which are derived from the same verb stem *kath-*. A cognate of the Romani form *xate* word may be found for instance in a form at-

---

4 Grassman's Law relates to the dissimilation of aspiration, especially in Old Indian and Old Greek. A root consisting of aspirate consonant – vowel – aspirate consonant loses the aspiration of the root-initial consonant. The aspiration comes up again when the aspiration of the second consonant is lost for morphological reasons (cf. Fox 1995: 258).



## 5. CONCLUSION

Even though the quotative is preserved in only a few dialects, its range of forms and meanings is so limited that there is no doubt that it goes back to one source. We propose that Northern Romani *xate* is of Indic origin. It is derived ultimately from the Old Indo Aryan verb *kath* (Hindi *kahnā*) meaning 'tell, relate, say', possibly with an interference from forms derived from Old Indo Aryan *khyā* 'to tell'. Its basic meaning is a quotative marker. An Indic origin is much more likely than an independent development, as the verb itself is not preserved in Romani, and quotative markers are clear areal phenomena in South Asia, otherwise rare in European languages (except in formal and written styles). In some Romani dialects the quotative particle changed its meaning to 'pretend' and 'allegedly'. Both are unsurprising extensions of the original meaning. Whether the *ince/iče* forms in Rumanian and Bulgarian Romani are cognate remains to be established.

If this etymology is correct, it is yet another grammatical Indicism in Romani. Almost all of the inflectional and derivational morphology, the demonstratives, almost all quantifiers, pronouns etc. are undoubtedly of Indic origin. The quotative is both formally and functionally very close to quotatives in modern Indic languages.

REFERENCES<sup>7</sup>

- |              |      |  |
|--------------|------|--|
| Allen, W.S.  | 1957 | Some phonological characteristics of Rajasthani, <i>BSOAS</i> 20: 5–11.  |
| Bakker, P.   | 1997 | Athematic morphology in Romani: The borrowing of a borrowing pattern, in: Matras, Y./Bakker, P./Kyuchukov, H. (eds.) <i>The Typology and dialectology of Romani</i> , Amsterdam: 1–21.   |
|              | 1999 | The Northern branch of Romani: mixed and non-mixed varieties, in: Halwachs, D.W./Menz, F. (eds.) <i>Die Sprache der Roma. Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext</i> , Klagenfurt: 172–209. |
| Bhatia, T.K. | 1987 | <i>A History of the Hindi grammatical tradition. Hindi-Hindustani grammar, grammarians, history and problems</i> , Leiden.   |
| Boretzky, N. | 1989 | Zum Interferenzverhalten des Romani (Verbreitete und ungewöhnliche Phänomene), <i>Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung</i> 42: 357–374.  |
|              | 1994 | <i>Romani. Grammatik des Kalderaš-Dialekts mit Texten und Glossar</i> , Wiesbaden.   |
| Bubenik, V.  | 1995 | On typological changes and structural borrowing in the history of European Romani, in: Matras, Y. (ed.) <i>Romani in Contact. The History, Structure and Sociology of a Language</i> , Amsterdam: 1–24.  |
| Burton, [R.] | 1849 | A grammar of the Játakí or Belochkí dialect. <i>Journal of the Bombay branch of the Royal Asiatic Society</i> 3(1): 81–125.  |

---

<sup>7</sup> A copy of Schultze's grammar was kindly made available to us by the Royal Library in Copenhagen. Ian Hancock provided us with the Jataki grammar and Anthony Grant sent us a copy of Winstedt's paper. Yaron Matras helped us with some references.



- Calvet, G. 1993 *Dictionnaire tsigane-français, dialecte kalderash*, Paris.
- Demeter, R.S./Demeter, P.S. 1990 *Cygansko-russkij i russko-cyganskij slovar' (kelderarskij dialekt) / Gypsy-English dictionary (Kalderash Dialect)* [Russian-Romani and Romani-Russian Dictionary] (ed. by L.N. Cherenkov), Moskva.
- Fox, A. 1995 *Linguistic Reconstruction. An Introduction to Theory and Method*, Oxford.
- Gilliat-Smith, B. 1927 Notes on Dr. J. Sampson 'the dialect of the Gypsies of Wales', *Journal of the Gypsy Lore Society* 6: 22–32.
- Hancock, I. 1998 *The Indian origin and westward migration of the Romani people*, Manchaca.
- Kāchru, Y. 1979 The quotative in South Asian languages, *South Asian Language Analysis* 1: 63–77.
- Mānuš, L. 1997a Čigānu-Latviešu-Angļu etimoloģiskā Vārdnīca [Romani-Latvian-English etymological dictionary], in: Mānuš, L. et al. *Čigānu-Latviešu-Angļu un Latviešu-Cigānu Vārdnīca*, Riga: 11–134.
- 1997b Īsa latviešu čigānu dialektu gramatika [Brief grammar of the Latvian Gypsy dialect], in: Mānuš, L. et al. *Čigānu-Latviešu-Angļu un Latviešu-Cigānu Vārdnīca*, Riga: 319–352.
- Masica, C.P. 1991 *The Indo-Aryan Languages*, Cambridge.
- Miklosich, F. 1878 Über die indische Heimat der Zigeuner and die Zeit der Auswanderung aus Indien. Pp. 287–296 (Part 3) of *Beiträge zur Kenntnis der Zigeunermundarten. IV*. Wien. Also in Leipzig 1984 reprint.
- Patkanov, K.P. 1911 Textes en Romani russe, *Journal of the Gypsy Lore Society* N.S. 4: 121–130.
- Rüdiger, J.C.C. 1990 *Von der Sprache und Herkunft der Zigeuner aus Indien. Leipzig 1782* (Nachdruck der Ausg. Leipzig, 1782 mit einer Einleitung von Harald Haarmann. *Linguarum minorum documenta historiographica*, 6), Hamburg: i–xxx, 37–84.
- Sampson, J. 1926 [1968] *The Dialect of the Gypsies of Wales*, Oxford.
- Schultze, B. 1745 *Viri plur. Reverendi Benjamin Schulzii Missionarii Evangelici Grammatica Hindostanica collectiis in diuturna inter Hindostanos Commoratione in justum Ordinem redactis ac larga Exemporum [sic] luce perfusis regulis constans et Missionariorum usui consecrata. Edidit et de suscipienda barbararum Linguarum Cultura præfatus est D.Jo.Henr. Callenberg*. Halæ Saxonum: in typographia Instituti judaici.
- Smart, B.C./Crofton, H.T. 1875 *The dialect of the English Gypsies*, London [reprint Detroit 1968].
- Turner, R. 1959 Transference of Aspiration in European Gypsy, *Bulletin of the School of Oriental and African Studies* 22. [Reprinted in *Collected papers 1912–1973*, London 1975].
- 1966 *Comparative dictionary of Indo-Aryan languages*, Oxford.
- Winstedt, E.O. 1944 Some Anglo-Romani etymologies, *JGLS*, 3rd series, 23/1-2: 67–68.

Peter Bakker

Institute for Linguistics, University of Aarhus

Heinz Werner Wessler

Department of Indic Philology, University of Aarhus