

回鹘文文献语言数词研究

阿卜杜外力·麦麦提江

中央民族大学中国少数民族语言文学学院，北京

收稿日期：2025年9月28日；录用日期：2025年10月28日；发布日期：2025年11月12日

摘要

本文基于历史比较语言学理论，考察了7~14世纪回鹘文献中的数词系统。通过对冯·加班(耿世民译，2003)、张铁山(2005)、马塞尔·埃尔达尔(刘钊译，2016)及陈宗振(2016)等经典研究的整合分析，全面梳理了回鹘文文献语言中的基数词、序数词、约数词、集合数词和分配数词五类范畴的形态特征、构词规则及句法功能。

关键词

回鹘文，数词研究，历史比较，数词范畴，形态学

A Study on Numerals in the Old Uyghur of Documents

Abudulwaili Maimaitijiang

School of Chinese Ethnic Minority Languages and Literatures, Minzu University of China, Beijing

Received: September 28, 2025; accepted: October 28, 2025; published: November 12, 2025

Abstract

This paper, based on the theory of historical comparative linguistics, investigates the numeral system in Old Uyghur documents from the 7th to the 14th centuries. Through an integrated analysis of classical studies such as those by Annemarie von Gabain (translated by Shimin Geng, 2003), Tieshan Zhang (2005), Marcel Erdal (translated by Zhao Liu, 2004), and Zongzhen Chen (2016), it comprehensively delineates the morphological characteristics, word-formation rules, and syntactic functions of five categories of numerals in the Old Uyghur language: cardinal numerals, ordinal numerals, approximative numerals, collective numerals, and distributive numerals.

Keywords

Old Uyghur, Numeral Study, Historical Comparison, Numeral Categories, Morphology

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

回鹘文作为维吾尔族历史上采用阿拉伯字母文字之前使用最广、传世文献较丰富的文字体系，承载着古代回鹘及周边突厥语部族的文明基因，在中亚与中国北方民族文化交流进程中占据核心地位。回鹘文源于中亚粟特文。最初，以七河流域为主要聚居地的突骑施部，率先采用粟特字母拼写突厥语。后因回鹘人对该文字的广泛接纳，渐被通称为“回鹘文”[1]。回鹘文的使用跨越宗教边界，佛教徒、摩尼教徒、景教徒及伊斯兰教徒均以其记录经典与日常；载体亦极为多元，涵盖碑铭、写本、木刻等。其影响更突破部族范围：元代为蒙古族采纳，经演变形成现代蒙古文(如《元史》载塔塔统阿以畏兀字(回鹘文)教蒙古贵族)；16世纪后，满族借蒙古字母体系发展出满文。在中亚，回鹘文还曾是金帐汗国、帖木耳帝国、察哈台汗国的官方文字(如“铁木耳库特鲁扎令” Temir Qutluq Yarliqi 等文献遗存)，新疆吐鲁番、哈密一带的使用延续至14~15世纪，明代的《高昌馆来文》可为实证[2]。

回鹘文文献堪称古代回鹘社会的“文化百科”，按内容可细分为四大类：1) 宗教文献：佛教类含《金光明最胜王经》《玄奘传》《俱舍论安慧实义疏》等；摩尼教类存《二宗经》《摩尼教徒忏悔词》；景教类见《福音书》《圣乔治殉难记》；伊斯兰教类含《帖木耳世系谱》《升天记》等。2) 文学作品：《弥勒会见记》《福乐智慧》(维也纳本)、《真理的入门》《乌古斯可汗传》等。3) 医学文献：如 Siddhasāra 等。4) 其他内容文献：《高昌馆来文》《高昌馆杂字》等大量社会经济文书、历法、两种语言对照的字典等。

回鹘文研究始于19世纪末20世纪初新疆南部、甘肃敦煌等地的文献发现，随即引发国际学术热潮，德国克拉普洛特(J. Klaproth)据《高昌译语》抄本研究古代维吾尔语文；米勒(F. W. K. Müller)、勒寇克(A. V. Le Coq)、班格(W. Bang)、冯加班(A. Von Gabain)与马塞尔·埃尔达尔(Marcel Erdal)等深化研究，刊布《回鹘文献研究》《高昌出土突厥语摩尼教文献研究》《古代突厥语语法》《古突厥语语法》等多卷成果涉及了回鹘文的数词研究，推动文献整理与语言语法分析。张铁山在《回鹘文献语言的结构与特点》中将回鹘文献数词系统分为基数词、序数词、集合数词、分配数词与约数词五类，详细阐述了其构成方式、语法特征及句法功能，为系统研究回鹘文数词奠定了重要基础[2]。李经纬在1990年发表的《回鹘文文献语言的数量词》中对回鹘文文献语言中的数词进行了分类，包括基数词、序数词、集合数词、分配数词、约数词和特殊数词，揭示了数词在文献中的多样性和复杂性[3]。此外，阿不里克木·亚森等在2001年发表的《吐鲁番回鹘文世俗文书语言数词研究》中以13~14世纪写成于吐鲁番地区的169件回鹘文世俗文书为语言资料，研究了回鹘文世俗文书语言中的数词，发现分配数词在世俗文书中出现频率很高，反映了当时社会生活中的实际需求[4]。张赫在2018年发表的《维吾尔语数词历时演变发展研究》中研究了维吾尔语数词的历时演变，指出数词的发展反映了语言在不同历史阶段的社会文化变迁[5]。综上所述，数词的研究不仅揭示了语言内部的规律，还展示了其与社会文化环境的互动关系，对于理解回鹘语及其背后的文化具有重要意义。

数词是表示事物数目和顺序的词类，在词汇、句法和形态上都具有鲜明的特征。根据其形式、词义和语法功能，回鹘语数词可系统性地分为基数词、序数词、约数词、集合数词和分配数词[2]。此外，基数词通过与其他词结合，可以表示倍数、分数、小数等概念，但这些并非独立的“倍数词”、“分数词”或“小数词”类别，而仅仅是相应的表示方法。纵观古今维吾尔语的发展历程，基数词、序数词、约数

词、集合数词的核心体系相对稳定，主要体现为语音层面的细微差异；相对而言，近代约数词的种类有所扩展。一个贯穿古今的重要语法特点是，数词普遍可以“名词化”，从而获得名词的复数、人称、格等形态变化[6]。

2. 基数词

基数词表示基本的、确切的数目，是构成其他复合数目和数词类别的基础。回鹘语的基数词体系包含单纯基数词和复合基数词。

2.1. 单纯基数词

一: bir	
二: eki/ekki	
三: üç	
四: tört	
五: beš/bäš	
六: altı	
七: yeti/yetti/yätti	
八: säkiz/säkkiz	
九: toquz	
十: on	
二十: yegirmi/ygirmi	
三十: otuz/ottuz	
四十: qırq/qırıq	
五十: älig	(此词与“手”同形，具有理据性)
六十: altmıš	(可能源自 altı “六”，但后缀-mıš 的构词理据尚不明确)
七十: yetmıš/yätmıš	(可能源自 yeti “七”)
八十: säkizon	(即 säkiz + on)
九十: toquzon	(即 toquz + on)
百: yüz	
千: bñq/bıq/mñq	
万: tümän	

万: tümän (普遍认为系借词，源自吐火罗语或中古波斯语[7]，除表确数“万”外，常引申表“众多”、“大量”之义，如 bñq yillıq tümän künlik bitig “千年万日的文书”，意即“永久的诏书”)。最大且无理据性的单纯基数词。

2.2. 复合基数词：由基本基数词组合而成

1) 十位与个位组合(11~19、21~29 等)：

回鹘语基数词最显著的特点之一体现在 11~19 以及 21~29 的表达上，采用“个位数 + 更高一级十位数”的“分段计数法”：

十一: bir yegirmi	(直译“一·二十”)
十二: eki yegirmi	(直译“二·二十”)

十三: üç yegirmi	(直译“三·二十”)
……以此类推至十九: toquz yegirmi	(直译“九·二十”)
二十一: bir otuz	(直译“一·三十”)
二十二: eki otuz	(直译“二·三十”)
二十三: üç otuz	(直译“三·三十”)
……以此类推至二十九: toquz otuz	(直译“九·三十”)

这种结构表明回鹘人将 11~20 视为一个计数段, 11 即为该段的第一个数, 故用 bir yegirmi (1+20) 表示。同理, 21~30 为下一段。这与现代以百年为单位的世纪划分法(如 1901~2000 为 20 世纪)有异曲同工之妙。

“余数”表示法(30 以上): 三十以上的数目, 常用“十位数 + artuqi (artuq ‘多余’ + -i 第三人称领属附加成分, 意为‘余数’) + 个位数”表示:

三十一: otuz artuqi bir
三十二: otuz artuqi eki
三十三: otuz artuqi üç
……以此类推至三十九: otuz artuqi toquz

此法有时也用于 40 以上的数字。

2) 直接组合法(40 以上):

40 以上的数字也常直接组合, 不严格遵循上述两法, 体现出一定的灵活性:

四十五: biş älig	(直译“五·五十”)
五十四: tört altmıš	(直译“四·六十”)
六十八: sākiz yetmiš	(直译“八·七十”)
七十二: eki sākizon	(直译“二·八十”)
八十五: biş toquzon	(直译“五·九十” 原文为 biş toquz on, 疑为 toquzon 的早期或变体形式)

3) 高位整数表示法:

“百”、“千”、“万”以上的整数, 采用“倍数词(1~9) + 单位词(yüz, bñ, tñmān)”的乘法关系构成:

二百: eki yüz
五百: biş yüz
八百: sākiz yüz
两千: eki bñ
四千: tört bñ
五千: biş bñ
七千: yeti bñ
一万: bir tñmān
四万: tört tñmān
五万: biş tñmān
……

其中, bir yüz (一百)、bir bñ (一千)开头的 bir 常省略。例如:

一百五十: <i>yüz älig</i>	(直译“百·五十”)
一百七十: <i>yüz yetmiş</i>	(直译“百·七十”)
一千二百: <i>bîñ eki yüz</i>	(直译“千·二百”)
一千五百: <i>bîñ biş yüz</i>	(直译“千·五百”)

4) 其他复合表示法:

① 高位 + *artuqi* + 低位: 此法更为通用, 不仅用于个位加十位, 也用于十位加百位等高位组合。如:

一百四十万: <i>yüz artuqi qırq tūmān</i>	(直译“百·多余·四十万”)
四百九十五: <i>tört yüz toquz on artuqi biş</i>	(直译“四百九十·多余·五”)
三十一岁: <i>otuz artuqi bir yaşıma</i>	(直译“三十·多余·一·岁”)

省去 *artuqi*: 晚期文献可见省略 *artuqi* 的形式, 如 *älig bir* (五十一)。

② 特殊高位词 *örki*: 表示百位以上的高位(源自动词 *ör-*“升起”), 如八百九十六: *säkiz yüz altı örki* (直译“八百·六·升起”)。值得注意的是, “九十九”可写作 *toquz örki* 或采用缺位表示法 *yüzkä bir ägsük* (直译“一百缺一”)。

③ 分数表示: 如 *iki ülügü atlıg ärti, bir ülügü yadag ärti* 通常译为“三分之二(字面: 它的两个部分)骑马, 三分之一(字面: 它的一个部分)步行”, 展示了早期表示分数的可能方式。

2.3. 基数词的历史演变与文献差异

回鹘文文献语言与更早的古代突厥文(如尼文)碑铭在基数词构成上的主要差异出现在约 13 世纪以后的晚期回鹘文文献中[3]。此时, 11~29 的表达方式常与早期文献不同, 趋向于更接近现代的直接组合法或“高位 + 低位”法:

十二: <i>on eki</i>	(而非早期的 <i>eki yegirmi</i>)
二十二: <i>yegirmi eki</i>	(而非早期的 <i>eki otuz</i>)
三十四: <i>otuz tört</i>	(而非早期的 <i>tört qırq</i>)
五十四: <i>älig tört</i>	(而非早期的 <i>tört altmış</i>)

此外, 回鹘文文献中还引入了外来数词, 如梵语的 *kolti* (千万)、*nayut* (那由它, 表无数), 汉语的 *ban* (万) [8], 以及复合词 *tūmān ban* (万万, 千千万万)。这些词多用于表示“大量”、“无数”的虚指意义, 而非精确计数。

表示“半”用 *yarım*, 如 *yarım* (一半), *iki yarım* (两个半)。在鄂尔浑和回鹘汗国碑铭中, *sıgar* 可能意为“一半”或“一部分”, 如 *sıgar süsi ... sıgar süsi ...* (一半军队……另一半军队……)。

3. 序数词

序数词表示事物的顺序或次第, 主要由基数词添加后缀 *-(X)nč* (变体包括 *-nč*、*-inč*、*-inč*、*-unč*、*-ünč*) 构成。例如:

第一: <i>birinč</i>
第二: <i>ikinč ~ ikinti</i>

值得注意的是, “第二”是唯一使用 *-nti/-ndi* 后缀的序数词形式, 其他序数词不使用此后缀。其来源可能与副词后缀 **+tl** 有关, 如 *amtı* “现在”。有一例 *äkin* 表示序数, 若非笔误则可能是 *äkinti* 的词基形

式[9]。

- 第三: üçünč
- 第四: törtünč
- 第五: bešinč
- 第六: altınč
- 第七: yetinč
- 第八: sākizinč
- 第九: toquzunč
- 第十: onunč

高位序数词遵循与基数词相同的构成逻辑:

- 第二十: yegirminč
- 第三十: otuzunč
- 第五十: äliginč
- 第七十: yetmişinč
- 第一万: tümäninč (此词在特定语境下有“末尾”、“最后”的自谦含义)

29 以上和以下的序数词构成法与基数词一致:

- bir yegirminč ay (第十一月)
- iki otuzunč bölük (第二十二章)
- qırqınč bölük (第四十章)

序数词 + (X)nč 后缀可能借自吐火罗语, 因其序数词后缀形式与此类似[9]。文献中也有 bir yegirminč (第十一)的用例。除规则派生外, “第一”常用其他词汇或短语表示, 如 başlayukı (由动词 başla- “开始”的副动词加形容词化后缀* + kI*构成)、baštinkı (“居首的”, 源自 baş “头” + 从格 + tIn + +kI)、aşnukı (如 äñ aşnukı kün “第一天”)。表示“第二十一以后”用 bir otuz küntä ken。

4. 约数词

约数词表示大约的、不确定的数目, 主要有以下几种构成方式:

基数词加-čä: 表示“大约……”、“……左右”。

- yüzčä (一百左右)
- äligčä (五十左右)

两个相邻基数词组合: 表示“……到……”、“几……”。

- iki üç (两三个)
- iki üç bñj (两三千)
- iki üç kün (两三天)
- biş on yıl (五到十年)

基数词与 artuq (多, 余)组合: 表示“……多”。

- yüz artuq (一百多)

üç yüz artuq	(三百多)
otuz artuq	(三十多)

5. 集合数词

集合数词表示由若干个体事物组合而成的整体概念，主要由基数词添加后缀-ağū/-ägū 构成。例如：

ikilägü ~ ikägü	(两个一起，俩)
üčägü	(三个一起，仨)
altağū	(六个一起)
yitägü	(七个一起)

部分学者认为后缀变体-ağū/-ägū (如 ikägūn、üčägūn)也用于构成集合数词。另有学者认为此处的-n是工具格附加成分。因为数词可以名物化并变格，所以 ikägūn 可分析为 ikägü + 工具格-n，意为“以两个一起的方式”；üčägūn 同理为 üčägü + -n，意为“以三个一起的方式”。birägü (由一个组成的一组)也有用例。

集合数词具有代词性(代表一组事物的名称)，在句中可接代词性宾格后缀 +nI 或领属后缀(此时带代词性 n)。当接工具格时，它们充当状语，表示行为的方式(“一起地”)。

与集合数词相关的还有：

İki + z (双胞胎)：此处的 + (X)z 是表示复数的成分(比较领属后缀 + (X)m + Xz “我们”，代词 siz “你们”)。

后缀 + gII 构成表示几何图形的名词，如 törtgil (四边形)。若 üçgil、törtgil、säkiz、kirlig、altı yegirmi kirlig 确指三、四、八和十六边形，则表明该后缀的适用性可能有限。

6. 分配数词

分配数词表示事物平均分配于各个个体的数目，由基数词添加后缀-ar/-är (用于辅音结尾的基数词后)或-rar/-rär (用于元音结尾的基数词后)构成。例如：

birär	(各一个)
ekirär	(各两个-eki 元音结尾 + -rär)
üčär	(各三个)
bišär	(各五个)
altırar	(各六个-altı 元音结尾 + -rar)
yetirär	(各七个)
onar	(各十个)
otuzar	(各三十个)
bišär yüz	(各五百，注意：仅第一个成分 biš 加分配后缀，yüz 不变)

分配数词通常在句中作定语，修饰被分配的名词：

onar ärkä	(给每十个人)
birär čingratgu	(各一个铃铛)

强调“仅各一”可用 birärkyä (birär + 表小/语用后缀 + kyA)。分配数词 birär 重叠(birär birär)或仅 bir 重叠(bir bir)可用作状语，表示“逐一”、“逐个地”：

birär birär adakın bap kāmīšip yünin kırkarlar	(他们逐个绑好羊腿，放倒在地，剪了羊毛)
birär birär kölmiş	(逐个地绑好)
bir bir aẓunlarda	(在每一世中，强调分配遍指)

在特定经济文书中，ikirär 有特殊用法，意为“各一半”或“各等份”，而非“各两个”：

...ikirär [ya]rīm bilišip [...] berür biz	(……我们各算一半……等量交纳)
...ik(i)rär t(ä)ñ bilišip t(ä)ñ berür biz	(……我们各算等份，等量交纳)

7. 数词的名词化

名词化是回鹘语数词的重要语法特征，使其能够像名词一样具有数、人称、格等形态变化，在句中充当名词性成分。文献中有大量实例。

7.1. 带格标记

ol ay biš yegirmigä käyrä başi ...	(当月十五日[与格]，在 käyrä 泉……)
taqıǵu yıl ikindī ay säkiz yegirmikä mana ilčigä ...	(鸡年二月十八日[与格]，我，伊里奇……)
bir otuz küntä ken	(二十一天以后[位格])

7.2. 带人称/领属标记

用于指称团体中的个体或部分。

ekisin özi aluuzdı. (他独自征服了两人[宾格 + 第三人称领属])。	
birisi uruq üzä mini boyunumın batı. (其中一个[主格 + 第三人称领属]夜叉用绳子拴住我的脖子)。	
bo üçägüdä birisin birisin titgülük idalaguluk kızıǵı kalsär ... (若轮到逐一放弃这三个[宾格 + 第三人称领属])。	
olarka üç ötmäk [berd]im. ol biri yılan tärkin [ka]pap yedi. (我给了他们三块面包。其中一条蛇[主格 + 第三人称领属]立即抢吃了)。	

特别值得注意的是基数词 bir (一)的名物化形式。按常规语法，以辅音 r 结尾，加第三人称领属附加成分应为+i，形成 biri。然而，实际文献中还出现了 birsi 的形式(可能是直接加-si 或 biri + -si 的叠加)。这种不符合常规的形式一直延续到现代维吾尔语(现代维吾尔语也有 biri 或 birsi)。

8. 其他相关表达

副词 öñi: 意为“独立于……，不同于……；独自地”。它通常支配位格或从格名词，但当与数词连用时，数词用主格形式，表达“分成……份/类”的特殊含义：

kop kamag yalñuklar üç öñi bölär.	(他将一切众生分为三类)。
yeti öñi ätözlärin bīčip.	(将她们碎尸七块)。

后置词 ara: 意为“在……之间”。当与 eki/iki (二)连用时，在鄂尔浑碑铭至晚期回鹘文献中均写作 äkin ara/ikin ara。后置词 ara 要求前面的词带宾格和领属后缀。这种形式可能暗示 äki/iki 的第二个元音 i 最初就是一个领属后缀，其词根 äk-可能与乌古斯语中表示“附加物”的词相关，故 äki 原义或为“其附加物”(指成对中的一个) [7]。

9. 结语

本文梳理了回鹘文文献语言中的数词体系，通过整合历史比较语言学视角与文献实证分析，揭示了

数词范畴的形态特征、构词规则及句法功能。回鹘文数词以十进制为基础,形成了基数词、序数词、约数词、集合数词和分配数词的完整体系,其构词规则既体现古代突厥语族语言的共性(如基数词分段计数法、序数词后缀-(X)nč的派生机制),又蕴含独特的文化逻辑。从历史演变看,回鹘文数词在保持核心体系稳定的同时,通过外来词汇吸纳(如梵语 kolti、汉语 ban)和构词创新(如特殊高位词 öрки、分配数词后缀 -rar/-rär)实现功能扩展,反映出古代回鹘社会与中亚、中原文明的互动。数词的“名词化”特征贯穿古今,使其兼具名词的形态变化(数、人称、格),这一语法现象不仅印证了回鹘语数词的灵活性,也为理解语言接触中的范畴扩展提供了典型案例。

参考文献

- [1] 耿世民. 古代维吾尔文献教程[M]. 北京: 民族出版社, 2006.
- [2] 张铁山. 回鹘文献语言的结构和特点[M]. 北京: 中央民族大学出版社, 2005.
- [3] 李经纬. 回鹘文文献语言的数量词[J]. 语言与翻译, 1990(4): 8-12.
- [4] 阿不里克木·亚森, 王正良. 吐鲁番回鹘文世俗文书语言数词研究[J]. 新疆大学学报(哲学社会科学版), 2001(1): 133-136.
- [5] 张赫. 维吾尔语数词历时演变发展研究[J]. 知识文库, 2018(16): 6, 8.
- [6] 陈宗振. 维吾尔语史研究[M]. 北京: 中国社会科学出版社, 2016.
- [7] Erdal, M. (1991) Old Turkic Word Formation: A Functional Approach to The Lexicon. Vol. I - II. Harrassowitz Verlag.
- [8] 冯·加班, 著. 古代突厥语语法[M]. 耿世民, 译. 呼和浩特: 内蒙古教育出版社, 2003.
- [9] 马塞尔·厄达尔. 古代突厥语语法[M]. 刘钊, 译. 北京: 民族出版社, 2016.