Rhee, Seongha. 2007. Different faces of equality: Grammaticalization of equative compatives in Korean. Japanese Korean Linguistics 15: 149-160.

The 15th Japanese-Korean Linguistics Conference was held at University of Wisconsin-Madison, from October 7- 9, 2005.

# Different Faces of Equality: Grammaticalization of Equative Comparatives in Korean*

SEONGHA RHEE
*Hankuk University of Foreign Studies*

## 1. Introduction

The grammatical notion of comparison seems to be among the conceptual primitives but its linguistic manifestations are diverse across many languages. This paper addresses the grammaticalization processes of 'equative comparatives' (Heine 1997), largely equivalent to English *like* and *as*, and their uses in Present Day Korean (PDK). It also presents interesting aspects that surfaced in the course of their grammaticalization, some conforming to the commonly attested grammaticalization principles while some deviating from them. Even though their interesting grammaticalization processes merit in-depth research, especially because there are multiple forms for the same or similar function, they have received little attention to date either individually or collectively, and this paper intends to fill the gap.

In PDK, there are six markers of equative comparative, i.e. *kathi, chelem, man(khum), taylo, tapkey* and *tus*[1] as shown in (1).[2]

(1) a. ku  ai-nun      appa-mankhum   khu-ta
    the  child-Top  father-EqCom  be.tall-Dec
    'The child is as tall as its father.'

  b. kunye-nun   ku-chelem   ttokttokha-ta
    she-Top     he-EqCom  be.smart-Dec
    'She is smart like him / as smart as he is.'

  c. kunye-nun   aki-kathi      kwiyep-ta
    she-Top     baby-EqCom  be.cute-Dec
    'She is cute like a baby / as cute as a baby is.'

  d. ney    mal-taylo      cham  coh-ta
    your   word-EqCom  very  be.good-Dec
    'It is good as you said.'

  e. ku-nun   senpi-tapkey      sal-ass-ta
    he-Top   scholar-EqCom  live-Pst-Dec
    'He lived as a (true) scholar.'

  f. pi-ka      o-l-tus                ha-ta
    rain-Nom come-Pros.Adn-EqCom  do-Dec
    'It looks like rain.'

## 2. Grammaticalization

Equative comparatives date back to the earliest Korean data in Hankul orthography, i.e. the 15th century, the Middle Korean (MidK) period, and they are still productively used in Modern Korean (ModK). In this section we investigate their lexical and structural sources.

---

[1] There are other markers of equative comparative in dialectal varieties, such as *meylo, maylo, maychilo, maynchilo, manki, tholem,* and numerous others (see Lee 2003), whose grammaticalization processes this paper does not address for interest of space.

[2] The transcription of the Korean data follows the Extended Yale Romanization System (Rhee 1996) to allow for old orthographic notations. The abbreviations used in gloss are: Acc: accusative; Adn: adnominal; Advz: adverbializer; Attm: attempt; Comp: complementizer; Conn: connective; Dec: declarative; Emph: emphatic; End: sentential ending; Eq.Com: equative comparative; Exclam: exclamative; Imp: imperative; Inst: instrumental; NF: non-finite; Nom: nominative; NP: noun phrase; Poss: possessive; Pres: present; Pros: prospective; Pst: past; Stat: status; and Top: topic.

## 2.1 Lexical and Structural Sources

The equative comparatives under discussion have diverse lexical sources as shown in Table 1. It is noteworthy that the majority, i.e. four out of six, have nominal sources, making reference to various aspects of a physical entity.

| Form | Lexical Source | Characteristics |
|------|----------------|-----------------|
| kathi | kath- 'be same' | adjective denoting identicalness |
| tapkey | tav- 'be qualified as' | defective adjective |
| chelem | thyey 'body' | Sino-Korean noun |
| man(khum) | man 'amount' | defective noun |
| taylo | tA 'place' | defective noun |
| tus | tus 'shape' | defective noun |

Table 1. Lexical Sources of Equative Comparatives

A comparison of their source structures shows a common feature that they largely make use of adverbializers such as *-lo, -i,* and *–key* as shown in Table 2. This has to do with the fact that adverbializer is a common device for phrasal or clausal connection in Korean.

| Form | Structural Source | Literal Meaning |
|------|-------------------|-----------------|
| kathi | kath + i 'ADVZ' | in the same manner |
| tapkey | tav +key 'ADVZ' | qualified as |
| chelem | thyey + lo 'INST/STAT' | with/as the body |
| man(khum) | man + (kom 'EMPH') | the exact amount |
| taylo | tA + lo 'INST/STAT' | with/as the place |
| tus | tAs | shape |

Table 2. Structural Sources of Equative Comparatives

## 2.2 Formal Change

The complexity in the historical development of the forms of each equative comparative is such that a comprehensive description of the developmental path is well beyond the scope of this paper. We will briefly describe their historical development.

Among the earliest forms is *kAthi,* an adverbialized adjective *kath-* 'be identical', which later developed into 'be similar' (Rhee 2004). There were numerous variants, and especially in the 18th century data diverse forms such as *kAschi, kAchi, kAtthi, kAsthi,* etc. are all attested. However, in ModK only *kathi* is used.

Another equative comparative with a long history is *tapkey,* a combination of a derivational suffix *-tav-* 'be qualified as, be equal to, be tantamount

to, be a good exemplar of, etc.' and the manner adverbializer -*key*. However, the further lexical origin of -*tav*- is unclear but it is hypothesized as 'be same' by some researchers (e.g. Kim 1998, Kim 1996, Huh 1975). The phoneme /v/ disappeared in ModK as a result of a general phonological change, and in all the words containing it, the phonemic value has changed into either /b/ or /w/. A variant form of the derivational suffix -*tAv*-, surviving in a very few fossilized lexical forms, is -*lAv*-. There was a change in the relative productivity of the two competing adverbializers -*key* and -*i,* the latter having been used more productively in early Korean (Rhee 1996). Thus the historically earlier form of *tapkey* is *tAvi*.

It is noteworthy that *chelem* was derived from a Sino-Korean noun *thyey* 'body', whose ModK counterpart is the palatalized *chey*. It first occurred in the mixed orthography of Chinese *thyey* and an instrumental -*lo* in the 16th century. It showed some variation between *thyelo* and *chyelo* in the 18th century, and only the palatalized variant survives in ModK.

The historical development of *man* and *mankhum* is complex. The ModK function of these two forms as equative comparative is to mark 'limit' and 'degree', and their historical origin goes back to *man/ma/maskam,* which meant 'amount' in the 15th century data. The attested variations are *man/mankom* in the 16th century, *makom* in the 17th century, *manchi* in the 18th century, and *mankhum* from the 19th century. A development worth noting is that from the 18th century, the productive particle *man* denoting 'only' began to be used as a result of split from *mankhum*. This *man* with the delimitation function and the *man* with the equative comparative function are in the relationship of homophony or polysemy (or even heterosemy, if we assign different grammatical categories to them), but we shall not go into the detail of this delimiter particle here.

*Taylo* is a combination of the defective noun *tA/ta* 'place' and the instrumental suffix -*lo*. It is attested in the 16th century in the form of *taylo* and its variant form *tAylo* is attested in the 18th century, which was more commonly used in the 19th century. However, from the 20th century only the first variant form, *taylo,* is used as a result of a general orthographic and phonological change in Korean that eliminated the phoneme /ɔ / of the vowel 'A', splitting it into three different vowels /a/, /ə/ and /ʊ/.

Finally, *tus* was originally a defective noun denoting 'shape' or 'state', occurring in the form of *tAs* in MidK, and often occurring with a light verb *ha*- 'do'. It has remained stable in form except for the general historical change that occurred to its vowel as noted above.

As is evident in the historical development of the equative comparatives in the preceding section, all these acquired the grammatical function, albeit through different paths and at different rates, from the MidK, i.e. the 15th through the 16th centuries.

**2.3 Semantic Change**

In general, semantic change is concomitant of grammaticalization and some-times it is even considered a prerequisite of grammaticalization. Interesting semantic changes are observable in the development of equative compara-tives.

As for *kathi,* the semantic change pattern can be characterized as in (2), an instance of semantic generalization gradually losing the degree of like-ness (Rhee 2004). The examples are as in (3).

(2) identicalness > similarity > non-dissimilarity
    identicalness > emphatic

(3) a. sinsa-kathi          yeyuypalukey   hayngtongha-yla
       gentleman-EqCom courteously     behave-Imp
       'Act courteously like a gentleman.'

    b. saypyek-kathi     talli-e-o-ass-ta
       dawn-EqCom     run-NF-come-Pst-Dec
       '(He) came very early in the morning.'

The semantic change of *tapkey* can be characterized as in (4) and its exam-ples are as given in (5).

(4) identicalness > similarity > qualification/membership

(5) a. pulkun  yen-s        koc-An       yenyenhi  kos-tap-tota
       red      lotus-Poss flower-Top  smoothly  flower-EqCom-Exclam
       'The red lotus flower is smoothly beautiful/fragrant.'
                                              (Twusienhay I: 7.2; 1481)

    b. ku-nun    tayhaksayng-tap-ta
       he-Top    college.student-Eq.Com-Dec
       'He is truly a college student (e.g. intelligent, etc.).'

The semantic change with *chelem* is an instance of generalization from iden-ticalness to similarity, in the sense that 'body' designates the extreme level of identicalness as does English *self.* The change from similarity to assumed similarity is an instance of subjectification. The semantic change pattern of *chelem* is as in (6) and some of the examples are given in (7).

(6)  body > shape > similarity > assumed similarity

(7) a. syonpatang-thyelo  stey-lAl     mAnt-Ala
       palm-EqCom        patch-Acc  make-Conn
       'Make  a  patch  in  the  shape  of  a  palm  (of  a  hand),  and
       then...'                              (Twuchangbang 29; 1663)

b. amwukes-to molu-nun    kes-chelem   hayngtongha-n-ta
   nothing-even not.know-Adn thing-EqCom behave-Pres-Dec
   '(He) behaves as if he did not know anything / pretends ignorance.'

As for *man* and *mankhum,* its development can be characterized as an instance of metaphor along the ontological continuum of [SPACE > TIME > QUALITY]. There is the subjectification phenomenon as well in the sense that some parts of the semantic change involve attribution of dynamicity, value, scalarity/relativity, and even causation, which are tied to the speaker's subjective judgments. The semantic change patterns are as in (8), and their partial uses are as exemplified in (9) and (10).

(8) man: amount > duration > degree > worth
    mankhum: amount > duration > degree > limit > cause/reason

(9) a. wupak-i    golf-kong-man   ha-ta
      hail-Nom   golf-ball-EqCom  do-Dec
      'The hail is as big as golf balls.'

   b. ku yenghwa-nun po-l-man       ha-ta
      the movie-Top   see-Adn-EqCom  do-Dec
      'The movie is worth watching.'

(10) a. mek-ul-mankhum   mek-ess-ta
        eat-Adn-EqCom    eat-Pst-Dec
       '(I) ate as much as I would like to eat.'

    b. cil-i       coh-un-mankhum     pissa-ta
       quality-Nom be.good-Adn-EqCom  be.expensive-Dec
       'It is expensive as much as /because it is good.'

The semantic development of *taylo* involves metaphorization from SPACE to QUALITY and CONCRETE to ABSTRACT as shown in (11). There is subjectification involved in the change as we can see attribution of dynamicity and scalarity/relativity. Some of the examples are given in (12).

(11) place > shape/quality/circumstance > fulfillment > conformity > exclusivity > emphatic

(12) a. yelay    kyesi-n    tAy-lAl    molA-zAvangita
        Buddha exist-Adn  place-Acc  not.know-End
       '(I) do not know the place where Buddha is.'
                                (Sekposangcel II, 11:10; 1495)

    b. iss-nun-taylo        noh-a-twu-ela
       exist-Adn-EqCom  leave-NF-keep-Imp
        'Leave it as it is.'

Finally, *tus* does not show any substantial semantic change. According to MidK data, especially in the Buddhist scripture commentaries, *tus* often corresponds to the Chinese *ye* 'similar' or *yuye* 'very similar', *sa* 'similar' or to the Korean *kAt-* 'be same/similar'. Since its syntactic environment warrants its analysis as a noun, albeit defective, we can assume the change as one from 'a shape having similarity' to 'in similarity to'. One aspect of the semantic development of *tus* is that it can be used for clausal connection equivalent to English *as if*, even though its meaning is ambiguous between factual and counterfactual descriptions.

## 3. Issues for Discussions

### 3.1 Layering and Specialization

If we look at the development of equative comparatives, their geneses are almost simultaneous despite the absence of evidence of mutual interaction in replacing their precedents in Old Korean. This is a state of affairs exhibiting an extreme case of 'layering' (Hopper 1991, Hopper & Traugott 2003). On the other hand, they show delicate semantic differences, and therefore, even though some of them are often interchangeable, the substitution often renders some level of semantic difference or of awkwardness. Therefore, there is a level of specialization according to their fine-grained subdivision of the functional domain, with a few members of the category being more dominant than others, a case of 'specialization' (Hopper 1991, Hopper & Traugott 2003). The specialization is organized along the variables of meaning and morpho-syntactic structure as shown in Table 3.[3]

| Form | Semantic Focus | Host | Frequency | Particle Ranking |
|------|----------------|------|-----------|------------------|
| kathi | similarity in shape | NP | 2,172 | 62/184 |
| takpkey | qualification | NP | N/A | N/A |
| chelem | (faked) similarity in quality | NP | 14,783 | 28/184 |
| mankhum | reaching standard | NP, Adn | 3,832 | 59/184 |
| taylo | establishing standard | NP, Adn | 3,540 | 48/184 |
| tus | shape/appearance | V, Adn | 13,780 | N/A |

Table 3. Specialization of Equative Comparatives

[3] The token frequency is based on the KAIST-KORTERM Corpus (13.6 million words), in which *tapkey* is not searchable as an independent linguistic unit. The particle ranking is based on National Academy of the Korean Language (2002), in which *tapkey* and *tus* are not classified as particles, and their rankings are thus not available.

As noted above, the functional specialization of these forms in subdivided domains is such that their substitution may show semantic differences as shown in (13) and (14).

(13)  a. kusalam-mankhum   mek-ess-ta
     b. kusalam-chelem       mek-ess-ta
      he-EqCom            eat-Pst-Dec
     a: '(I) ate as much as he did.' [amount]
     b: '(I) ate as he did.'        [appearance]

(14) a. ton-i          ku-mankhum  iss-umyen  ne-n       pwuca-ta
     b. ton-i          ku-taylo     iss-umyen  ne-n       pwuca-ta
      money-Nom that-EqCom   exist-if     you-Top  rich.person-Dec
     a: 'If you have that much money, you are rich.' [fulfillment of standard]
     b: 'If the money is still there, you are rich.'     [established standard]

The difference in the examples in (13) is due to the fact that *mankhum* focuses on the referenced entity's reaching the quantitative standard, whereas *chelem* focuses on the similarity or identicalness in quantity or state. Likewise, the difference in (14) is due to the fact that *taylo,* unlike *mankhum,* focuses on the established standard. This type of subtlety is present in most of the equative comparatives.


### 3.2 Universal Paths and Grammaticalization Channels

The grams of equative comparatives came from diverse source lexemes of different conceptual and grammatical domains as discussed in 2.1. However, all these forms, through sense attraction by virtue of conceptual affinity, proceeded to the grammaticalization of a single grammatical function. This state of affairs lends support to the grammaticalization principle called 'the Universal Path' as proposed by Bybee et al. (1994) or 'Grammaticalization Channels' as proposed by Heine and Reh (1984).

Indeed, in a brief survey through grammaticalization lexicons and elsewhere we see that there are diverse, yet similar, source lexemes as in (15).

(15) a. German *wie* 'how?' > *wie* 'like'
     b. Spanish *como* 'how?' > *como* 'like'
     c. Mandarin Chinese *you* 'exist' > *you* 'like, as ... as'
     d. Twi *sɛ* 'resemble' > *sɛ* 'as ... as' (Lord 1993)
     e. Tamil *poola* 'be similar with' > *poola* 'like, as'  (Lehmann 1989)
     f. Idoma *ka* 'say' > *kɛ*  'like' (Lord 1993)
     g. English OE *all-swá* 'wholly so' > *as* (OED)
     h. English OE *líc* [OTeut 'body' > *like* (OED)

i. Thai *yàaŋ* 'way, manner' > *yàaŋ-kàb* ['way/manner-with] 'as if' (Bi-sang 1998)
j. Kenya Pidgin Swahili *namna* 'manner' > *namna* 'like, as' (Heine & Kuteva 2002)

In view of the development of Korean equative comparatives and those attested in other languages, we can construct a hypothetical path followed by the equative comparatives in the course of their grammaticalization as in Figure 1.
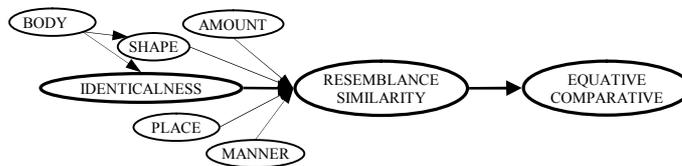


Figure 1. Grammaticalization Paths of Equative Comparatives

### 3.3 Borrowings in Source Lexemes

Borrowed lexemes are less likely to grammaticalize into grammatical markers since they are often associated with specific genres and styles and consequently a low use frequency (Rhee 1996). The low frequency of Sino-Korean words in contrast with native Korean words is well illustrated in the following state of affairs of Sino-Korean words in the Korean lexis.

(16)  About 55-70% in the entire ModK vocabulary
      About 50% in the top 1,000 frequent content words
      About 5% in the top 100 frequent morphemes
      Only one in the 800 particles/endings, i.e. *chelem*

As shown in (16) the fact that *chelem,* despite its Sino-Korean origin, was grammaticalized is noteworthy. As a matter of fact, the lexical origin has become completely opaque in ModK due to its extensive grammaticalization processes. Another factor that contributed to the resultant opacity is a rare and seemingly unmotivated phonological addition, i.e. from /če/ to /čərəm/, contra the Parallel Reduction Hypothesis (Bybee et al. 1994), according to which grams lose semantic content and phonological volume in parallel in the course of grammaticalization.

### 3.4 Source Characteristics

The next issue for discussion is the source characteristics of the equative comparatives. From the morphosyntactic point of view, their co-occurrence patterns follow our expectation that they would retain the characteristics of their source constructions (the 'Persistence Principle'; Hopper 1991, Hopper & Traugott 2003). For instance, those equative comparatives of nominal sources may allow for NP-hosts as instances of compounding, or Adnominal-hosts as instances of relativization. This is well observed by *mankhum* and *taylo*. Similarly, *kathi* of an adjectival source does not allow for Adnominal-hosts, which is also expected from the Persistence Principle.

However, there are instances that are not in consonance with this expectation. For example, *chelem,* which has a nominal source, does not allow for an Adnominal-host. In ModK all adnominal phrases or clauses must accompany a semantically vacuous noun *kes* 'thing' to host the equative comparative *chelem.* Likewise, *tus,* also having a nominal source, cannot take an NP-host, and can take finite or non-finite verb as its host, both of which are against the expectation. We might suspect their original idiosyncratic behavior as being responsible, but a full explanation as to these anomalies should await a further investigation.

Another aspect of source characteristics is the phonological shapes. Among the most common principles relating to the phonology of grammaticalizing forms is Formal Erosion (Lehmann 1995[1982], Heine et al. 1991, Bybee et al. 1994). The equative comparatives *taylo, kathi, tus,* and *tapkey* did not undergo much phonological change other than following other general phonological changes that occurred in the history of Korean. There are no reductive processes involved. Though this does not strictly conform to the Parallel Reduction Principle or Formal Erosion, the degree of violation by these forms may be said to be minimal. However, *mankhum* and *chelem* have undergone additive processes, a sheer violation of the Parallel Reduction Principle or Formal Erosion.

Still another aspect of the source characteristics relates to their semantics. The six markers of equative comparative, in a situation of conflict of interest because of their shared grammatical function, exhibit delicate division of labor by each specializing in more fine-grained grammatical subfunctions resorting to such notions as scalarity, identicalness, similarity, conformity, concomitance, etc. in contemporary Korean. Their current semantics may be characterized in the following terms. *Mankhum* still has the 'amount' meaning inherited from the source. *Chelem* does not have the 'body' meaning, and has undergone the most extensive semantic bleaching. Consequently it is used most frequently. *Kathi* still has the 'be same/similar' meaning and retains the source meaning most transparently. Possibly attributable

to its formal and semantic transparency, it is used least frequently. *Taylo* still has the 'standard/background' meaning inherited from the source. *Tus* still has the 'shape' meaning (or more frequently the 'appearance' meaning), its function extending to include factual and counterfactual description of appearance, and it is used in high frequency. *Tapkey* still retains the original 'qualification' or 'qualified membership' meaning, thus undergoing minimal semantic change. Since it is predominantly used for an adjectival derivation by hosting a noun and as many words derived from such a process have been fossilized as independent lexical items, its use is highly restricted.

All these descriptions of current semantic characteristics point to the facts that most of them retain, though in varying degrees, the original semantics of the source lexemes, and that the speed of semantic change is by no means uniform.

## 4. Conclusion

In this paper we have looked at the six equative comparatives used in ModK. Drawing upon historical data from the MidK literature and a contemporary corpus, this paper attempted to trace and contrast the developmental paths of these equative comparatives, focusing on their genesis with respect to lexical and structural sources, on their conflict of interest in the same functional domain, which was resolved by specializing in subdivided functions, and on their observance and/or violation of grammaticalization principles widely subscribed to in grammaticalization scholarship, such as universal path, persistence, parallel reduction and formal erosion. The diversity presented by the grammaticalization history of these grammatical forms shows different faces hidden behind the cover label of equality.

## References

Bisang, W. 1998. Adverbiality: The view from the Far East. In J. van der Auwera (ed.) *Adverbial Constructions in the Languages of Europe.* Berlin: Mouton de Gruyter. 643-812.

Bybee, J. L., R. D. Perkins, & W. Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago: Chicago University Press.

Heine, B. 1997. *Cognitive Foundations of Grammar.* Oxford: Oxford University Press.

Heine, B., U. Claudi, & F. Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework.* Chicago: University of Chicago Press.

Heine, B., & T. Kuteva. 2002. *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Heine, B., & M. Reh. 1984. *Grammaticalization and Reanalysis in African Languages.* Hamburg: Helmut Buske.

Hopper, P. J. 1991. On some principles of grammaticalization. In E. C. Traugott & B. Heine (eds.) *Approaches to Grammaticalization*. 2 vols. Vol. 1: 17-35. Amsterdam & Philadelphia: John Benjamins.

Hopper, P. J. & E. C. Traugott. 2003[1993]. *Grammaticalization*. Cambridge: Cambridge University Press.

Huh, W. 1975. *Wuli Yeysmalpon: 15-seyki Kwuke Hyengthaylon* [Old Korean Grammar: 15th Century Korean Morphology]. Seoul: Saem Publishing.

KAIST KORTERM Corpus, http://morph.kaist.ac.kr/kcp/ accessed September 2005.

Kim, J.-A. 1998. *Cwungseykwukeuy Pikyokwumwun Yenkwu* [A Study of Comparative Constructions in Middle Korean]. Seoul: Thaehaksa.

Kim, C. S. 1996. *Kwukeuy tanehyengsengkwa tanekwuco yenkwu* [A study of Korean word formation and word structure]. Seoul: Thaehaksa.

Lee, K. 2003. *Kwuke Pangen Mwunpep* [Grammar of Korean Dialects]. Seoul: Thaehaksa.

Lehmann, C. 1995[1982]. *Thoughts on Grammaticalization*. Newcastle: LINCOM Europa.

Lehmann, T. 1989. *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

Lord, C. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam/Philadelphia: John Benjamins.

National Academy of the Korean Language. 2002. *Hyentay Kwuke Sayong Pinto Cosa: Hankwuke haksupyong ehwi sencengul wihan kicho cosa* [A Use Frequency Survey of Modern Korean: A pilot survey for selection of Korean educational vocabulary]. Seoul.

Rhee, S. 1996. Semantics of Verbs and Grammaticalization: The Development in Korean from a Cross-Linguistic Perspective. Doctoral dissertation, The University of Texas at Austin. Seoul: Hankuk Publisher.

Rhee, S. 2004. How far likeness can go: Grammaticalization of *kath-* in Korean. Paper presented at Berkeley Linguistics Society 30. February 13-16, 2004.