

THE SYNTAX OF PRAGMATICALIZATION*

Jila Ghomeshi
University of Manitoba

Persian has SOV word order when direct objects are nominal and SVO order when verbs take clausal complements. It is a null subject language. This paper is a preliminary attempt to bring together the uses of the multifunctional item *ke* under a theory of syntactic change.

1. The uses of *ke*

In this section I outline the various uses and functions of *ke* starting with the most grammatical and ending with the least.

1.1 Complementizer for relative clauses

From a descriptive point of view, *ke* is probably best characterized as an element that introduces subordinate clauses of various types. In formal syntax it thus belongs to the category Complementizer. One such use of *ke* is to introduce a relative clause, as in (1a). Relative clauses in Persian can be, and frequently are, extraposed, meaning that they are separated from the noun they modify and appear clause-finally as in (1b). Given that *ke* ‘moves’ with the relative clause, we see that it forms a constituent with the clause that follows and not with the preceding noun.

- (1) a. un mænzel [CP-Rel **ke** dust=dasht-im] -o foruxt-æn
that house **that** friend=have+1PL +OM sold+ 3PL
‘They sold that house that we liked.’
- b. un mænzel-o foruxt-æn [CP-Rel **ke** dust=dasht-im]
that house+OM sold+ 3PL **that** friend=have+1PL
‘They sold that house that we liked.’

Estaji (2011) proposes that this use of *ke* may have developed from a relative pronoun *hya* (Old Persian) via *kē* (*čē*), (*ī*) (Middle Persian). The path from relative pronoun to complementizer is documented within the

*Many thanks to Khashayar Hamidzadeh and Sharareh Esmaeili who worked with me as research assistants in the spring of 2013, and thanks to the Department of Linguistics, University of Manitoba, for making this possible. Additional thanks to Sharareh who took on the topic of the functions of *ke* for a course paper before knowing what she was getting into. I would like to acknowledge the students in my Grammatical Polysemy class (Winter 2012) whose enthusiastic discussion and individual research projects were inspiring. Thanks also to the audience at the 2013 CLA in Victoria for stimulating questions and discussion. All errors are my own.

grammaticalization literature (see for example the discussion of English *that* from pronoun to complementizer in Hopper & Traugott (2003:190-4) and for a more recent change that is still in progress see Brook (2011) on the use of the locative relative pronoun *where* as a relative complementizer).

1.2 Complementizer for adverbial subordinate clauses

In Persian there are many expressions that serve to subordinate one clause to another. They are often themselves complex and prepositional in origin, though they act as chunks. They introduce adjuncts and the resulting clauses are understood to have temporal, concessive, conditional, result, or purpose meanings (to name a few). Examples are given in (2) but see also Lazard 1957/1992:236-254, and Perry (2007:2.8.4-2.8.7).

- (2) a. bæ'd æz in **ke** 'after' (lit. after that which)
 b. ta **ke** 'so that'
 c. bæra-ye in **ke** 'for, because' (lit. for that which)
 d. chun **ke** 'because'
 e. ægær **ke** 'although'

Of significance for this paper is that almost all such expressions end with *ke*, which is cliticized to the preceding element. According to Estaji (2011), this use of *ke* may have developed from the connective *yadi*, *yaθa*, *hyat*, *yaθa* (Old Persian) via *ka* (Middle Persian).

1.3 Complementizer for complement clauses

The prototypical use of *ke* as a complementizer is with verbs of saying or knowing where it introduces a selected complement and is not a reduced version of a more complex expression. In these cases it is optional:¹

- (3) a. *introducing direct discourse* (Perry 2007)
 goft [CP (**ke**) mæn ne-miy-am]
 said.3SG **that** I NEG+come+1SG
 'He said "I'm not coming."'
 b. *introducing indirect discourse* (Perry 2007)
 goft [CP (**ke**) ne-miy-ad]
 said.3SG **that** NEG+come+3SG
 'He said he's not coming.'
 c. *introducing an indicative complement clause*
 mi-dun-æm [CP (**ke**) aftab daq-e]
 DUR+know+1SG **that** sun hot+3SG
 'I know (that) the sun is hot.'

¹ The following abbreviations are used in the interlinear glosses: DUR= durative, OM= object marker, EZ= ezafe (a linker), SG= singular, PL= plural, SBJN= subjunctive, FOC= focus marker, NEG= negation prefix.

Estaji (2011) states that this use of *ke* may have developed from the complementizer *tyat* (Old Persian) via *ku* (Middle Persian).

While it is evident from the examples above that *ke* has a complementizer function, assuming that it has *only* this function leads to a bi-clausal analysis of examples like those in (4) below. In these examples we see that *ke* can introduce the subjunctive complement to a raising verb (4a), the subjunctive complement of a control verb (4b), and the subjunctive complement to a modal verb (4c):

- (4) a. *introducing the subjunctive complement clause of a raising verb*
 lazem-e [(ke) in mænzel-o be-frush-im]
 necessary+3SG that this house+OM SBJN+sell+1PL
 ‘It is necessary that we sell this house.’
- b. *introducing the subjunctive complement of a control verb*
 mi-xa-m [(ke) in mænzel-o be-frush-æm]
 DUR+want+1SG that this house+OM SBJN+sell+1SG
 ‘I want to sell this house.’
- c. *introducing the complement of a modal verb*
 bayæd [(ke) in mænzel-o be-frush-æm]
 must that this house+OM SBJN+sell+1SG
 ‘I must sell this house.’

Ghomeshi (2001) argues that control verbs such as *xastæn* ‘want’ in (4b) and modal verbs such as *bayæd* ‘must’ in (4c) take vP, not CP, complements in Persian. This means *ke* must be something other than a complementizer in these constructions, a point that is acknowledged but not very satisfactorily addressed in Ghomeshi (2001:54, see also references cited therein). In fact, the appearance of *ke* in examples such as these may be best analyzed as instances of its use as a modal particle which is outlined in the next section.

1.4 *Ke* as a modal particle

Consider the following uses of *ke*, which have been characterized as “idiomatic” (Lazard 1957/1992) or emphatic (Windfuhr 1979:70):

- (5) a. A: xob hæva chetor-e?
 well/so weather how+be.3SG
 ‘So what’s the weather like?’
- B: hæva bæd nist, xub-e
 weather bad NEG.be.3SG, good+be.3SG
- inja ke xeyli xub-e
 here PRT very good+be.3SG
 ‘The weather’s not bad; it’s nice; here, it’s really nice.’

- b. jai ke ne-mi-r-in emshæb?
 place PRT NEG+DUR+go+2PL tonight
 ‘You’re not going anywhere tonight, are you?’ (presumed answer is ‘no’)
- c. shoma ke nahar mehmun-in
 you PRT lunch guest+be.2PL
- mæn-o mizar-in xune-ye nayyer
 1SG+OM put+2PL home+EZ Nayyer
 ‘You guys are invited over to lunch, so you’ll drop me off at Nayyer’s.’

[above examples from CALLFRIEND Farsi FA_4099, Canavan & Zipperlen 1996]

Lazard (1992:255) observes that this construction “is extremely frequent in colloquial language: one (or sometimes several) of the noun phrases of a sentence ... is found at the beginning and followed by *ke*, after which comes the rest of the clause. This construction usually has the effect of emphasizing the term...” The identification of this use of *ke* as an emphatic element meaning something like ‘voilà, mais, eh bien’ was made as early as Chodzko (1852:141, cited in Windfuhr 1979:70).

In the examples above *ke* appears in second position, i.e. after the first constituent of the clause. As Lazard notes, this is a feature of the colloquial or spoken language. However *ke* is not limited to this position and as others have noted it may appear on any constituent. Oroji & Rezaei (2013)² give the following examples:³

- (6) a. mæn-ke ketab-o be Ali ne-mi-d-æm
 I+PRT book+OM to Ali NEG+DUR+give+1SG
 ‘I won’t give the book to Ali.’
- b. mæn ketab-o-ke be Ali ne-mi-d-æm
 I book+OM+PRT to Ali NEG+DUR+give+1SG
 ‘I won’t give THE BOOK to Ali.’
- c. mæn ketab-o be Ali-ke ne-mi-d-æm
 I book+OM to Ali+PRT NEG+DUR+give+1SG
 ‘I won’t give the book to ALL.’

² I would like to thank Negin Ilkhanipour for bringing this article to my attention.

³ The examples have been adapted to conform to the transcription and glossing system used in this paper.

- d. mæn ketab-o be Ali ne-mi-d-æm-ke
 I book+OM to Ali NEG+DUR+give+1SG+PRT
 ‘I won’t GIVE the book to Ali.’

Notably absent from this paradigm is an example of *ke* in an embedded clause.⁴ Such examples can be constructed but are hard to find in naturally occurring data (e.g. in the CALLFRIEND corpus, Canavan & Zipperlen 1996). It is unclear whether this is because they are possible but rarely produced or because they are not possible at all. I leave this issue for further research.

The appearance of *ke* clause-finally, seen in (6d) above, highlights a similarity between *ke* and sentence-final particles in general. For instance, among the functions Bateni (2010) identifies as associated with the particle use of *ke* is one he terms ‘authentication for the listener’. This use is illustrated in the following example, provided by Bateni:

- (7) shoma charshænbe be xane-ye pærviz mi-ay-id-ke?
 you Wednesday to house+EZ Parviz DUR+come+2PL+PRT
 ‘You’re coming to Parviz’s house on Wednesday, aren’t you?’

The translation into English of *ke* as a tag accords with its function of checking or confirming information with the addressee. Lazard (1992:257) states that *ke* “gives to the sentences an interrogative value calling for an affirmative answer and, in exclamations, underlines the obvious.” Building on an example that Lazard (1992:255) provides in this regard, we can get the following contrasts:

- (8) a. qæhve mi-xor-id?
 coffee DUR+consume+1PL
 ‘Do you drink coffee.’
- b. qæhve-ke mi-xor-id?
 coffee+PRT DUR+consume+1PL
 ‘You drink coffee (at least)?’ [scalar reading]
- c. qæhve mi-xor-id-ke?
 coffee DUR+consume+1PL+PRT
 ‘You drink coffee, don’t you?’

Oroji & Rezaei (2013) propose that *ke* is a focus particle, however they also assert that information structure does not play a role in what can be focused. That is, “any constituent whether old or new from the viewpoints of both hearer and discourse, can be focalized.” (Oroji & Rezaei 2013:82) Moreover, their corpus study reveals that *ke* most often appears on subjects (50 out of 122 instances) and in terms of parts of speech most often appears on pronouns (35 out of 122 instances). Pronominal subjects in a null subject language such as

⁴ I would like to thank Martina Wiltschko for posing a question about this, which is worthy of further exploration.

Persian are usually considered to be added for ‘emphasis’, not focus. For these reasons, it seems implausible to analyze *ke* as a focus particle without stretching the meaning of focus to something very vague. Indeed, the readings shown above are almost impossible to unite into one concise meaning whether we call that meaning focus or not. I will therefore simply call *ke* a modal particle, the properties of which will be discussed in the next section. Apart from the issues around meaning I will also turn to the question of how such particles evolve *syntactically* over time.

2. Modal Particles and Pragmaticalization

Within the substantial body of literature on grammaticalization (e.g. Lehmann 1985, Hopper 1991, Hopper & Traugott 2003, Heine 2003, Narrog & Heine 2011), certain types of semantic change have been identified. Perhaps the best known is desemanticization, i.e. the phenomenon of semantic ‘bleaching’ or reduction, that occurs as lexical items take on grammatical uses (Heine 2003:579). Desemanticization well describes the semantic change that verbs undergo when they develop into auxiliaries. However, not all meaning change involves loss. Traugott (1989:35.5) identifies a change whereby meanings that encode or externalize speaker perspectives and attitudes are acquired by a lexical item, which she calls subjectification.

Subjectification in grammaticalization is the development of a grammatically identifiable expression of speaker belief or speaker attitude to what is said. Like the original hypothesis [...] subjectification is not limited to grammaticalization but can also be found in lexical change, for example, in such well-known cases of pejoration as *boor* ‘countryman, farmer’ > ‘crude person’. (Traugott 2003:633-643, as quoted in Diewald 2011:373)

Traugott (1989) shows how subjectification accompanies the syntactic and morphosyntactic changes that verbs undergo to become first deontic and then epistemic modals.

In the two examples discussed thus far, the expressions undergoing change remain part of the grammar (as auxiliary or modal verbs). However, some changes result in expressions that are seen as outside the grammar. In these cases the endpoint of change is a pragmatic or discourse function, hence the term ‘pragmaticalization’ (Erman & Kostinas 1993, Aijmer 1997). Whether pragmaticalization is an independent process or ought to be subsumed under grammaticalization as argued, for example, by Traugott (2007) and Diewald (2011) is not the concern of this paper. Either way the process that is named by the term ‘pragmaticalization’ is an apt characterization of the use of *ke* as a particle that is under consideration here.

There are two properties that have been purported to distinguish pragmaticalization from grammaticalization: (a) pragmaticalized elements express a pragmatic or discourse function and (b) they are outside the grammar. Both of these properties require further elaboration. The first property will be discussed in this section and the second property in the section that follows.

Traugott (2007, see also references cited therein), in her summary of three papers on discourse markers, discusses the ways in which discourse markers differ from modal particles. She states (pp. 140-1) that discourse markers sequence units of talk; they point forward or backward in the discourse. Examples include Italian *allora* ‘then’ and Spanish *bien* ‘well’. Modal particles, in contrast, lack connective properties at the discourse level. She notes (p. 145) that the contexts in which they occur are dialogic, and they often mark assertions in contexts of counter-expectation, i.e. are ‘adversative’. Examples include *denn, doch* in German, and *bien, donc* in French. In German their occurrence in the “Middle-field” is considered important, though they can also appear in other clause-internal positions. Phonologically, modal particles lack stress, and semantically have inferential, epistemic meanings (p. 142). Finally she notes that modal particles are often untranslatable from one language to another and are often deletable in translation.

By these criteria it is clear that *ke* falls more into the category of modal particle than discourse marker. It is not used to sequence units of talk but rather in adversative contexts. It occurs most often in second position in the clause. Moreover, the range of meanings associated with its use as a particle are hard to translate succinctly.

In the next section of this paper we turn to the syntax of the modal particle *ke* in an effort to understand how it has developed from its complementizer use.⁵

3. The syntax of grammaticalization vs. pragmaticalization

Grammaticalization, as it refers to the changes that a lexical item undergoes over time, is commonly associated with the development of inflectional morphology. For instance, the following grammaticalization cline from Hopper & Traugott (1993:7) represents the way in which an independent content word can take on grammatical uses and change into a bound morpheme:

(9) content item > grammatical word > clitic > inflectional affix

This cline shows the way in which a pronoun, for example, can be reanalyzed as a clitic and then as an agreement affix.

Van Gelderen (2011, and earlier work cited therein) seeks to formalize these kinds of changes within the Minimalist Program (Chomsky 1995, et seq). With reference to the cline in (9), she proposes that there is a principle of Feature Economy that strips away semantic and interpretable features, leaving only uninterpretable features:

⁵ While I know of no diachronic study to support the claim that the particle use of *ke* is the newer one, I believe it is a valid assumption based on the fact that it is a feature of spoken rather than written discourse.

- (10) **Feature Economy**
 Minimize the semantic and interpretable features in the derivation, for example:
 Adjunct Specifier Head Affix
 semantic > [iF] > [uF] > [uF]
 [van Gelderen 2011:14.17]

Continuing with the example of the way in which pronouns develop into agreement affixes, the Principle of Feature Economy explains this in terms of features. A pronoun loses its ability to refer as an independent element (loss of semantic and interpretable features) and becomes a bound element that enters into relationships with other constituents in the clause by virtue of having uninterpretable features that must be checked.

Van Gelderen (2011) provides two other principles that account for the changes lexical items undergo as they become grammaticalized. The Head Preference Principle posits that lexical items that occur in phrasal positions (e.g. as complements or specifiers) will come to be reanalyzed as heads.

- (11) **Head Preference Principle (HPP)**
 Be a head, rather than a phrase.
 [van Gelderen 2004, 2011:13.15]

This principle explains, for example, the tendency for relative pronouns (merged as specifiers within CP) to be reanalyzed as complementizers (merged as C-heads).

The third principle involved in syntactic change is what van Gelderen terms the Late Merge Principle (see also Roberts & Roussou 1999, 2003):

- (12) **Last Merge Principle**
 Merge as late as possible.
 [van Gelderen 2011:14.17]

This is consistent with Chomsky's (1995, 2001) 'merge-over-move' principle according to which, rather than merging an element lower in a syntactic structure and then moving it higher, the element is merged in the higher position. This principle has been invoked to explain the change from main verb to auxiliary. Perhaps more relevantly, this tendency for elements to merge at higher and higher points in a syntactic structure accompanies semantic changes such as subjectification, discussed in the previous section.

If we turn back to *ke*, it is fairly clear that the syntactic principles outlined thus far do not help explain how its use as a modal particle is related to its use as a complementizer. It has not taken on properties that make it look like a piece of inflectional morphology as Feature Economy would predict. While the Head Preference Principle explains its path from relative pronoun to complementizer (see section 1.1), this principle does not relate the complementizer use to the particle use, as they are both head-like. Finally, as a complementizer *ke* is arguably merged as the highest head in a syntactic structure so there is nowhere else for it to go under the Late Merge Principle. In section 2, however, we noted that a defining property of pragmaticalized

elements is that they seem to be ‘outside the grammar’. Let us consider what this might mean.

I propose that the syntax of pragmaticalization involves the loss of selectional features – those that are involved in determining head-complement relations. This means that heads cease to be constituent-forming elements at all. Let us call this the principle of Detachment:

- (13) **Detachment Principle**
 [head complement] > [_{adjunct} head ~~complement~~]

In the case of *ke*, it goes from being a complementizer that heads a CP and takes a TP complement, to a category-less particle that is (phonologically) enclitic and that can appear anywhere within a clause:

- (14) [_{CP} *ke* [_{TP} ...]] > [_{PRT} *ke* [~~TP~~ ...]]

Note that examination of *ke* in this paper has led to the claim that the syntax of pragmaticalization involves the bundling of three properties: (a) a lexical item loses its ability to take a complement, (b) a lexical item becomes category-less, and (c) a lexical item becomes syntactically mobile. These properties have been identified in the change in English of “complement taking mental predicates” such as *I think*, *I suppose*, into adverbial-type expressions with epistemic meanings (see Van Bogaert 2011, for instance). It is possible, however, that these three changes might not always occur simultaneously – an issue that merits further investigation.

As a final point, I would like to suggest that the idea that modal particles are ‘outside the grammar’ means they are not part of syntax proper. They are not elements that undergo the operation Merge and participate in feature checking. This points to the existence of a post-syntactic level of composition at which constituents can be “framed” for pragmatic purposes such as exclamation, protest, sarcasm or humour. The division of labour between syntax proper and this level is an exciting area for future exploration.

References

- Aijmer, Karin. 1997. *I think* – an English modal particle. In *Modality in Germanic languages. Historical and comparative perspectives*, eds. Toril Swan and Olaf Jansen Westvik, 1-47. Berlin & New York: Mouton de Gruyter.
- Batani, Mohammad Reza. 2010. Use of “*ke*, *dige*, *axe*, *ha*” in spoken Persian. *Bokhara journal*. [Translation provided by Sharareh Esmaeili]
- Brook, Marisa. 2011. One of those situations where a relative pronoun becomes a complementizer: a case of grammaticalization in action ... again. In *Proceedings of the 2011 annual conference of the Canadian Linguistic Association*, ed. Lisa Armstrong.
- Canavan, Alexandra and George Zipperlen. 1996. CALLFRIEND Farsi LDC96S50. CD-ROM. Philadelphia: Linguistic Data Consortium.
- Chodzko, Alexander. 1852. *Grammaire persane ou principes de l'iranien moderne, accompagnées de fac-similés pour servir de modèle d'écriture et de style pour la correspondance diplomatique et familière*. Paris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

- Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1-52. Cambridge: MIT Press.
- Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49: 365-390.
- Erman, Britt and Ulla-Britt Kostinas. 1993. Pragmaticalization: The case of *ba'* and *you know*. *Studier I Moderspråkvetenskap* 10: 76-93.
- Estaji, Azam. 2011. [1389]. A historical study of homophonous *ke* in Persian. *Journal of Linguistics and dialects of Khorasan* 2: 1-13. [Translation provided by Sharareh Esmaeili]
- Gelderen, Elly van. 2004. *Grammaticalization as Economy*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Gelderen, Elly van. 2011. *The Linguistic Cycle*. New York: Oxford University Press.
- Ghomeshi, Jila. 2001. Control and Thematic Agreement. *Canadian Journal of Linguistics* 46:9-40.
- Heine, Bernd. 2003. Grammaticalization. In *The handbook of historical linguistics*, eds. Brian Joseph and Richard Janda, 575-601. Oxford:Blackwell.
- Hopper, Paul J. 1991. On some principles of grammaticalization. In *Approaches to Grammaticalization*, eds. Elizabeth Closs Traugott and Bernd Heine, 17-35. Amsterdam & Philadelphia: John Benjamins.
- Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Hopper, Paul J. and Elizabeth Closs Traugott. 2003. *Grammaticalization* (2nd ed.). Cambridge, New York: Cambridge University Press.
- Lazard, Gilbert. 1957. *Grammaire du Persan contemporain*. Paris: Klincksieck.
- Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. English translation. Costa Meca, CA: Mazda Publishers.
- Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20: 303-318.
- Narrog, Heiko and Bernd Heine. 2011. *The Oxford Handbook of Grammaticalization*. Oxford, New York: Oxford University Press.
- Oroji, Mohammad Reza, and Amir Rezaei. 2013. Exploring 'ke' as a Focus Particle in Persian from both Form and Function Points of View. *Australian Journal of Linguistics* 33: 76-84.
- Perry, John R. 2007. Persian Morphology. In *Morphologies of Asia and Africa*, ed. Alan S. Kaye, 975-1019. Winona Lake, Ind.: Eisenbrauns.
- Roberts, Ian and Anna Roussou. 1999. A Formal Approach to "Grammaticalization". *Linguistics* 37: 1011-1041.
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65, 31-55.
- Traugott, Elizabeth Closs. 2003. Constructions in grammaticalization. In *The handbook of historical linguistics*, eds. Brian Joseph and Richard Janda, 624-647. Oxford:Blackwell.
- Traugott, Elizabeth Closs. 2007. Discussion article: Discourse markers, modal particles, and contrastive analysis, synchronic and diachronic. *Catalan Journal of Linguistics*, 6: 139-157.
- Van Bogaert, Julie. 2011. *I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics* 49:295-332.
- Windfuhr, Gernot L. 1979. *Persian Grammar: History and State of its Study*. The Hague: Mouton Publishers.