Morphologization in Turkish: Implications for Phonology in Grammaticalization[1]

René Schiering (University of Leipzig)

1. Introduction

Grammaticalization theory assumes that the gradual progression from a content item to a grammatical marker is accompanied by a number of interdependent phonological, morphosyntactic, and functional processes. Accordingly, morphologization processes, such as cliticization and compounding, are said to be concomitant with phonological erosion and desemantization (Lehmann [1982] 1995, Heine & Reh 1984, Heine, Claudi & Hünnemeyer 1991, Hopper & Traugott 1993, Croft 2003). Some proponents of this theory even claim that the loss of autonomy and substance defines grammaticalization as opposed to other mechanisms in language change, for instance reanalysis (Haspelmath 1998).

The role of phonology in grammaticalization has recently been reconsidered in the context of a cross-linguistic study on cliticization (Schiering 2006). A number of phonological processes, namely structure preservation, assimilation, weakening and strengthening, can accompany ongoing grammaticalization. The distribution of these phonological rules can be predicted by a rhythm-based typology of language which distinguishes between mora-, syllable- and stress-based languages. With respect to erosion, mora- and syllable-based languages tend to retain the phonological substance of cliticized elements which ultimately leads to disyllabic clitics and affixes. Stress-based languages, on the other hand, tend to reduce and delete

the phonological substance of cliticized elements developing subminimal clitics and affixes in the course of morphologization.

This paper will test the predictions made by the rhythm-based typology against diachronic data from Turkish. In Section 2, Turkish will be situated in the rhythm-based typology of language and the latter's predictions concerning the phonology of grammaticalization will be summarized. In Section 3, two morphologization processes in the verbal domain, namely the cliticization of pronouns yielding agreement markers and the univerbation of verb stems resulting in new tense-aspect morphology, will be traced from Old Turkic to Modern Turkish. The phonological changes involved in these processes will be of focal interest. Finally, the actual diachronic data from Turkish will be compared with the predictions made by the rhythm-based typology and the major findings will be discussed in the context of grammaticalization theory.

2. Turkish in a Rhythm-based Typology of Language

The conception of linguistic rhythm which will be adhered to in this study is based on the assumption that different phonological properties tend to cluster in the prototypes of mora-, syllable- and stress-based rhythm (Auer 2001, Dufter 2003, Schiering 2006). In what follows, we will restrict our attention to the phonetic correlates of stress, the segmental effects of stress, the degree of syllable complexity, the distribution of length contrasts and the domain for vowel harmony and their respective distribution over the rhythmic prototypes.

Stress may be realized phonetically by pitch only or by a combination of pitch, duration and intensity (Beckman 1986). Phonetically weak stress is prototypical for mora- or syllable-based rhythm, whereas stress-based rhythm is characterized by phonetically strong stress. In Turkish, the realization of stress relies on pitch movement and to a certain extent on intensity.[2] With respect to this parameter, Turkish behaves like a mora- or syllable-based language.

---

[2] All information and data on Old Turkic and Modern Turkish have been taken from the following reference grammars: Gabain (1950), Erdal (2004), Kornfilt (1997) and Lewis (2000).

Stress-based languages further exhibit strong segmental effects of stress, i.e. vowel reduction and consonant weakening in unstressed syllables and vowel lengthening and consonant strengthening in stressed syllables (Bybee et al. 1998). Although there are some hints at word-medial vowel alternation and deletion, e.g. *ayïz > aɣzan* 'being said' (Old Turkic) and *burada > burda* 'here' (Modern Turkish), Turkish lacks such segmental effects of stress and behaves like other mora- or syllable-based languages with respect to this parameter.

Another phonological property which distinguishes stress-based from mora- and syllable-based languages is the degree of syllable complexity (Blevins 1995). Languages of the former class show high degrees of syllable complexity, whereas languages of the other types have simple or moderate syllable structure. Since Turkish allows only six syllable types, it belongs to the latter group of languages, cf. *o* 'he/she/it', *at* 'horse', *bu* 'this', *sol* 'left', *ilk* 'beginning', *kırk* 'forty'.

Mora- and syllable-based languages behave alike with respect to the phonological parameters discussed in the preceding paragraphs. The distribution of length contrasts in vowels and consonants, however, helps to distinguish the two rhythmic types. Mora-based rhythm is characterized by length contrasts with are distributed irrespective of stress placement. Although phonemic vowel length can be reconstructed for the proto-language, Turkish exhibits vowel length only in loans, *beraber* 'together'. Long vowels surface phonetically after contractions, e.g. *değil* /deyil/ > [di:l] 'not', and in expressive lengthening, e.g. *asla > aslaaa* 'never!'. Geminates occur only underlyingly in loans such as *hak* 'the right, justice' and word-medially at morpheme boundaries of complex words, such as *bat-tı* 'it sank'. Accordingly, Turkish cannot be considered a mora-based language.

Although the rhythm-based typology makes reliable predictions with respect to the prosodic and phonotactic parameters discussed above, the distribution of morphophonological rules, such as cluster simplification and coalescence, turns out to be erratic over the rhythmic prototypes. Vowel harmony, however, is restricted to disyllabic domains in stress-based languages and applies over word domains in languages of other rhythm classes. In Turkish, palatal and

labial assimilation are not restricted to smaller domains, but span the word domain, e.g. *anla-yacak* 's/he will understand' and *göz-lük-çü* 'optician, oculist'.

The distribution of the various phonological properties over the rhythm prototypes is illustrated in Figure 1. Note that Turkish behaves like a typical representative of syllable-based rhythm. In fact, the rhythmic profile of Turkish has not changed in the course of the development from Old Turkic to Modern Turkish.

Figure 1: Selected phonological properties of rhythmic prototypes

	Mora-based	*Syllable-based*	*Stress-based*
Accent	none/weak	none/weak	strong
Stress effect	none/weak	none/weak	strong
Syllable types	simple	moderate	complex
Length contrasts	unrestricted	restricted	restricted
Vowel harmony	word	word	disyllabic

The rhythm-based typology of language makes a number of predictions with respect to possible phonological effects of grammaticalization. Within stress prosodies, such as Turkish, the model predicts gradual stress reduction, i.e. from primary stress to secondary stress to unstressed, and gradual integration into the word domain for stress placement, i.e. from unstressable to stressable (see Selkirk 1995 for a formal analysis). Whereas this stress reduction goes hand in hand with segmental reduction in stress-based rhythm, languages of the other rhythm classes have no reductive potential in their phonologies and retain the grammaticalized element. Accordingly, we would not expect erosion to accompany grammaticalization in a syllable-based language like Turkish. Since vowel harmony processes are sensitive to the word domain in languages of this rhythmic type, gradual integration into the word domain should also manifest itself in the inclusion of the grammaticalized element into the vowel harmony domain. Since morphophonological rules operating at the morpheme boundary occur in languages of all rhythmic classes, junctural processes such as cluster simplification and coalescence are possible and provide the

only context in which segments can be lost in the course of prosodic integration.

3. Morphologization in Turkish

To test the predictions made by the typology outlined in the previous section, we will examine two cases of grammaticalization in Turkish, namely the cliticization of personal pronouns which led to subject agreement marking (Givón 1976) and the univerbation of verbal complexes which led to new aspect-tense markers (Lehmann [1982] 1995).

In Old Turkic, sentences with pronominal subjects were formed with a postponed pronoun at the end of the non-verbal or verbal predicate. For the sake of emphasis or contrast, another personal pronoun could be placed in preverbal subject position. In thirteenth century texts, the postponed pronouns appear cliticized to the preceding word. In Modern Turkish, the cliticized subject pronouns form the back bone of the *z*-paradigm of subject agreement marking (cf. Adamović 1985, Kornfilt 1996, Good & Yu 2005). The various stages in this diachronic development are exemplified for the first person singular in the examples (1)-(3).

(1) *(ben) kelür ben* 'I am coming'

(2) *(ben) gelür-ven ~ (ben) gelür-em* 'I am coming'

(3) *(ben) gelír-im* 'I come'

Although the construction as such did not change significantly, the phonological status of the bound morphemes underwent a number of changes. Assuming that both *kelür* and *ben* constitute words in (1), the combination is prosodized as a phonological phrase in which the first word receives stress. In (2) and (3), this phrasal stress has been reinterpreted as irregular word stress in which the prominence lies on the last syllable of the host (cf. Kabak & Vogel 2001 and Inkelas & Orghun 2003). Thus, the cliticized element does not get prosodically integrated in the domain for word-final stress. The reduction of stress from phrasal to word-level stress is not accompanied by segmental reduction but by the integration of the cliticized element into the

vowel harmony domain. Additionally, we encounter the application of *b > v* sandhi at the host-clitic boundary. These changes are responsible for the change in the surface realization of the cliticized element, from the invariant *ben* to the harmonizing *van ~ ven* with assimilated initial consonant. Noteworthy, other changes in the form of the marker are not due to phonological rules but to morphological shifts based on analogies. The second variant *(y)am ~ (y)em* which surfaces as *-em* in example (2) is a verb ending which has been copied from the optative and analogically extended to the aorist. Both exponents for the first person singular had been lost by the fifteenth century. The morpheme which nowadays expresses this category in the *z*-paradigm evolved from the contamination with the possessive suffix *(y)um ~ (y)üm* in the fifteenth century. After this marker underwent regular delabialization, it now surfaces with fourfold vowel harmony.

Although the actual development of the various morphemes which participated in the paradigm over the course of time is rich in detail, the generalization that the phonological processes involved hardly reduce the segmental composition holds true. Typically, these processes are restricted to assimilation rules, e.g. devoicing *-dur > -tur* '3.sg.' and labialization *-viz > -vüz* '1.pl.', and cases of regular sound change, e.g. *siŋiz > siniz* '2.pl.'. There is only one case in which cluster simplification leads to the loss of a segment, i.e. the loss of *v* in *Türk-vüz > Türk-üz* 'we are Turks'. However, more effective are the morphological shifts which are an essential part of the development of each individual marker.

Another morphologization process which can be traced throughout the documented history of Turkish has its origin in the converb construction which consists of a non-finite verb marked by a converb marker and a finite verb marked for aspect, tense, mood, person and number (Heine & Kuteva 2002). The examples in (4)-(6) illustrate Old Turkic converb constructions which have been grammaticalized to express actionality, intention, ability and version (cf. Erdal 1979, Johanson 1998, Ağcagül 2004).

(4) *geli yür-* 'to be coming'   (*yüri-* 'to go')

(5) *kör-ü bil-* 'to know how to obey'   (*bil-* 'to know')

(6) *alta-yu tur-* 'to keep cheating'   (*tur-* 'to stand')

Although the converb marker changed in some of the cases, the construction itself is still in use in Modern Turkish. With respect to their grammatical status, the various forms vary in such a way that they can be interpreted as representing different stages of morphologization on a grammaticalization cline. At least *-Iyor* 'progressive' (7) and *-(y)Abil* 'potential' (8) can now be considered disyllabic suffixes, since they are completely desemanticized and form an uninterruptible coherent and cohesive morphological word with their base.[3]

(7) *gel-iyor-um* 'I am coming'   (*yor-* 'to go')

(8) *gel-ébil-ir-im* 'I can come'   (*bil-* 'to know')

(9) *söylen-é-dur-ur* 'he keeps grumbling'   (*dur-* 'to stand')

Unlike the cliticization phenomena discussed above, the univerbation of these verbal complexes is accompanied by very little phonological effects. Assuming again that the source construction consisted of two phonological words which received regular phrasal stress on the first word, stress has been reduced and reinterpreted as irregular word stress on the first vowel of the disyllabic suffix. Prosodic integration into the domain for word-final stress does not apply and in contrast to cliticization compounding is not paralleled by the prosodic integration into the vowel harmony domain. Note that the second vowel of both *-Iyor* and *-(y)Abil* is invariant. Apart from some sound changes which occurred, the only noteworthy phonological process which applied in the development of the progressive suffix is haplology, which has been a regular process in Old Turkic, cf. *sür-ür-či > sürči* 'painter', *yür-ür > yür* 's/he goes', *dur-ur > dur* 's/he stands'. It is important to emphasize that the loss of the second syllable is not due to ongoing grammaticalization, but to the regular application of the phonological rule of haplology at that diachronic stage of the language.

---

3 Additional evidence for the different degrees of grammaticalization comes from the possibility of inserting the clitic *=dA* between the converb marker and the second verb stem. This seems to be possible with *-A-dur-*, less acceptable with *-Abil-* and presumably ungrammatical with *-Iyor-* (see also Bainbridge 1988).

4. Summary and Conclusion

In the preceding section we discussed the phonological effects of two morphologization processes in Turkish, cliticization of pronouns and univerbation of verbal complexes. As predicted by the prosodic cline for stress phonologies, we encounter stress reduction from phrasal to word-level stress in both cases. However, stress placement has not been altered in the discussed constructions. The synchronic irregularity of stress assignment in morphologically complex forms which evolved in morphologization is thus due to the lack of prosodic integration. The reduction of stress is not accompanied by segmental reduction in the cases we discussed above. What we find in cliticization is prosodic integration into the vowel harmony domain and the application of junctural sandhi processes, e.g. assimilation and cluster simplification. In this context, the rhythm-based typology proves a reliable model for predicting possible segmental effects of grammaticalization. The evidence from univerbation, on the other side, has to be taken with a grain of salt, since in this case ongoing grammaticalization is not accompanied by integration into the vowel harmony domain. As a result, the morphologically complex forms which evolved in morphologization form irregular domains with respect to vowel harmony. Note that the elements in question are subject to regular sound change to the same extent as elements which do not undergo grammaticalization. In the evolution of the *z*-paradigm of subject agreement, morphological shifts which are based on analogy form a substantial ingredient of the diachronic development and are inseparably intermingled with the grammaticalization process.

The diachronic evidence compiled in this paper provides overall positive evidence for the rhythm-based typology of phonology in grammaticalization. The syllable-based rhythm of Turkish prohibits the erosion of grammaticalized elements which ultimately leads to the accretion of morphological markers. The prosodic clines which have been proposed as part of the rhythm-based typology, however, do not necessarily mirror diachronic change. Although stress reduction occurs, the gradual integration of grammaticalized elements into the stress domains of their hosts has never been a factor in the development of the forms discussed above. As already demonstrated

in other studies, for instance Kabak & Schiering 2004, different grammaticalization processes may trigger different phonological processes even in a single language. Whereas cliticization in Turkish leads to the spreading of vowel harmony and the application of junctural processes, univerbation is not accompanied by such processes. Given the fact that, even in the diachrony of morphologization in a single language, grammaticalization cannot be characterized in a uniform manner, the findings of this study cast doubt on universal scenarios for grammaticalization and its independence from other forms of morphological change, such as analogy and reanalysis, as expressed in 'grammaticalization theory.'

References

Adamović, Milan 1985. *Konjugationsgeschichte der türkischen Sprache*. Leiden: Brill.

Ağcagül, Sevgi 2004. Grammaticalization of Turkic postverbial construction. *Orientalia Suecana* 53, 5-14.

Auer, Peter 2001. Silben- und akzentzählende Sprachen. In: Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.) *Language typology and language universals*. Berlin, New York: de Gruyter. 1391-1399.

Bainbridge, Margaret 1988. -(y)E-bil- and the conjunction dE: splitting the unsplittable. In: Koç, Sabri (ed.) *Studies in Turkish linguistics*. Ankara: METU. 95-115.

Beckman, Mary E. 1986. *Stress and non-stress accent*. (Netherland phonetic archives 7.) Dordrecht: Foris.

Blevins, Juliette 1995. The syllable in phonological theory. In: Goldsmith, John A. (ed.) *The handbook of phonological theory*. (Blackwell handbooks in linguistics, 1). Oxford: Blackwell. 206-244.

Bybee, Joan & Chakraborti, Paromita & Jung, Dagmar & Scheibman, Joanne 1998. Prosody and segmental effect. Some paths of evolution for word stress. *Studies in Language* 22, 267-314.

Croft, William 2003². *Typology and universals*. (Cambridge textbooks in linguistics.) Cambridge: Cambridge University Press.

Dufter, Andreas 2003. *Typen sprachrhythmischer Konturbildung*. (Linguistische Arbeiten 475.) Tübingen: Niemeyer.

Erdal, Marcel 1979. Die Konverb- und Aoristendungen des Alttürkischen. *Ural-Altaische Jahrbücher* 51, 104–126.

Erdal, Marcel 2004. *A grammar of Old Turkic*. (Handbook of Oriental studies, Section 8, Central Asia 3.) Leiden: Brill.

Gabain, A. von 1950². *Alttürkische Grammatik*. (Porta linguarum orientalium 23.) Leipzig: Harrassowitz.

Givón, Talmy 1976. Topic, pronoun, and grammatical agreement. In: Li, Charles N. (ed.) *Subject and topic*. New York: Academic Press. 149-188.

Good, Jeff & Yu, Alan C. L. 2005. Morphosyntax of two Turkish subject pronominal paradigms. In: Heggie, Lorie & Ordóñez, Francisco (eds.) *Clitic and affix combinations*. Amsterdam: Benjamins. 315-341.

Haspelmath, Martin 1998. Does grammaticalization need reanalysis? *Studies in Language* 22, 315-351.

Heine, Bernd & Claudi, Ulrike & Hünnemeyer, Friederike 1991. *Grammaticalization. A conceptual framework*. Chicago, London: University of Chicago Press.

Heine, Bernd & Kuteva, Tania 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Reh, Mechthild 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Hopper, Paul J. & Traugott, Elizabeth C. 1993. *Grammaticalization*. (Cambridge textbooks in linguistics.) Cambridge: Cambridge University Press.

Inkelas, Sharon & Orgun, Cemil O. 2003. Turkish stress: a review. *Phonology* 20, 139-161.

Johanson, Lars 1998. The structure of Turkic. In: Johanson, Lars & Csató, Éva Á. (eds.) *The Turkic languages*. London, New York: Routledge. 30-66.

Kabak, Barış & Vogel, Irene 2001. The phonological word and stress assignment in Turkish. *Phonology* 18, 315-360.

Kabak, Barış & Schiering, René 2004. A corpus study on the distribution of Turkish *ise* and its clitic form. *Turkic Languages* 8, 232-244.

Kornfilt, Jaklin 1996. On copular clitic forms in Turkish. *ZAS Papers in Linguistics* 6, 96-114.

Kornfilt, Jaklin 1997. *Turkish*. (Descriptive grammars.) London, New York: Routledge.

Lehmann, Christian [1982] 1995. *Thoughts on grammaticalization*. (Lincom studies in theoretical linguistics 1.) München: Lincom Europa.

Lewis, Geoffrey 2000². *Turkish Grammar*. Oxford: Oxford University Press.

Schiering, René 2006. *Cliticization and the evolution of morphology. A cross-linguistic study on phonology in grammaticalization*. Konstanz: Konstanzer Online-Publikations-System, Bibliothek der Universität Konstanz. [http://www.ub.uni-konstanz.de/kops/volltexte/2006/1872/]

Selkirk, Elisabeth O. 1995. The prosodic structure of function words. In: Beckman, Jill N. & Dickey, Laura W. & Urbanczyk, Suzanne (eds.) *Papers in optimality theory*. (University of Massachusetts Occasional Papers 18.) Amherst: GLSA. 439-469.