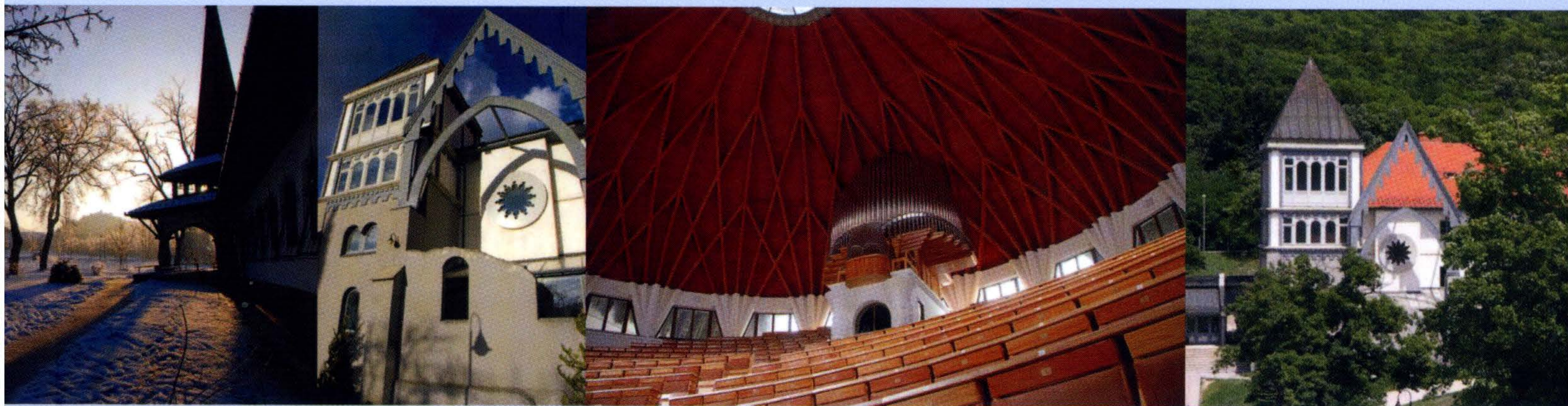


CONGRESSUS XI
INTERNATIONALIS
FENNO-UGRISTARUM

Piliscsaba 2010

Pars IV



FU11
2010

Congressus XI. Internationalis Fenno-Ugristarum
Piliscsaba, 9–14. VIII. 2010

Pars IV

Dissertationes sectionum: Linguistica



Piliscsaba, 2011

THE ADESSIVE OCCURRING WITH THE MAST-FORM:
A 20TH-CENTURY INNOVATION IN ESTONIAN?

1. According to the most authoritative descriptions of Standard Estonian (*eesti ühiskeel*), verbs of ‘allowing’ and ‘forbidding’ may occur in one or more of the four constructions with the *da*-infinitive (henceforth: **DA**) or the *mast*-form (henceforth: **MAST**). Let us illustrate the four constructions using the verb *keelama* ‘to forbid’ as a verb occurring in all of them:

- | | | | | | | |
|--------|------|--------|--------|-------|-----------|-------------------------------|
| (1) a. | Mees | keelas | sõbral | külla | tulemast. | (Adessive+MAST construction) |
| b. | Mees | keelas | sõpra | külla | tulemast. | (Partitive+MAST construction) |
| c. | Mees | keelas | sõbra | külla | tulla. | (Adessive+DA construction) |
| d. | Mees | keelas | sõpra | külla | tulla. | (Partitive+DA construction) |

The sentences in (1) are taken from Rätsep (1978: 196, 200). For the purpose of the present study, we assume that sentences (1a) – (1d) are all synonymous, meaning roughly ‘the man forbade a friend from coming for a visit.’ In this paper, ‘infinitive construction’ (IC) is the term reserved for the designation of each of the four types of sentences defined here.

The Ade+DA construction (*lasen tal midagi teha*) was already in use in the Estonian literary language of the 1890s, but the Ade+MAST construction (*keelan tal midagi tegemast*) appears to be of relatively recent origin, its inception dating back to the 1920s at the earliest. It seems that the ICs with the partitive – *lasen teda midagi teha* (Part+DA) and *keelan teda midagi tegemast* (Part+MAST) – used to be the standard ways of ‘allowing’ and ‘forbidding’ in Estonian.

The basic assumption for the present study can be stated as follows: A change in the grammar of a language is not a simple event describable as ‘X changed into Y at such-and-such a time’, but rather it is a composite process in which a number of factors interact among themselves over a certain period of time. We are going to show in this paper that an appropriately selected set of corpora can be an efficient source of data for the substantiation and dating of syntactic changes in a language as well as for the description of how various factors interact with each other in a particular language change.

2. In order to obtain an overall picture of the ICs in Standard Estonian, the following authoritative works have been consulted. These works will be referred to collectively as ‘Standard Grammar’ of Estonian in this paper. Table 1 summarizes the description given by Standard Grammar of the distribution of ICs.

Eesti keele grammatika I. Morfoloogia. Sõnamoodustus. Tallinn. 1995.

Eesti keele grammatika II. Süntaks. Lisa: Kiri. Tallinn. 1993.

Rätsep, Huno: *Eesti keele lihtlausete tüübid.* Tallinn. 1978.

Tauli, Valter: *Eesti grammatika II. Lauseõpetus.* Uppsala. 1980.

Tauli, Valter: *Standard Estonian grammar. Part II. Syntax.* Uppsala. 1983.

TABLE 1. Distribution of the ICs according to Standard Grammar

		MAST		DA	
		ADE	PART	ADE	PART
keelama	<i>forbid</i>				
takistama	<i>hinder</i>				
segama	<i>disturb</i>				
hoidma	<i>hold</i>				
hoiatama	<i>warn</i>				
pidurdama	<i>restrain</i>				
võõrutama	<i>alienate</i>				
aitama	<i>help</i>				
käskima	<i>command</i>				
laskma	<i>allow</i>				
lubama	<i>permit</i>				
paluma	<i>beg</i>				
soovitama	<i>suggest</i>				
võimaldama	<i>enable</i>				

Each white cell in Table 1 indicates that at least one of the works of Standard Grammar states explicitly that the verb in the leftmost column may occur with that particular infinitive construction. Otherwise cells are dark-colored.

3. Two corpora (1990s and 1920s corpora) are used to put forward a hypothesis, or more exactly a set of hypotheses. In addition, five smaller corpora, referred to collectively as ‘Historical corpus’ in this paper, are used for the validation of the hypothesis.²³

1990s corpus or *Tasakaalus korpus* (Balanced Corpus): a total of 15 million words consisting of 5 million words each of newspaper, literary and scientific texts from the late 1990s.

1920s corpus or *Asutawa Kogu protokollid* (Minutes of the Constitutional Assembly): a total of 1.93 million words of text from 1919–1920.

Historical corpus: a subset of *Eesti Kirjakeele Korpus 1890–1990* (the Corpus of Literary Estonian), consisting of five subcorpora (1890s, 1900s, 1910s, 1930s and 1950s; 43–117 thousand words each) containing newspaper articles and literary texts.

4. Tables 2 and 3 summarize the data from which the hypothesis is derived. The color-coding is the same as in Table 1 above. The asterisk indicates that the frequency counts given for the verb are estimations based on random sampling.

²³The three corpora are all available from the website of the Research Group of Computer Linguistics, University of Tartu:

Tasakaalus korpus:

<http://www.cl.ut.ee/korpused/grammatikakorpus/>

Asutawa Kogu protokollid:

<http://www.cl.ut.ee/korpused/baaskorpus/akp/>

Eesti Kirjakeele Korpus 1890–1990:

<http://www.cl.ut.ee/korpused/baaskorpus/>

TABLE 2. The distribution of infinitive constructions in the 1990s corpus

		MAST		DA	
		ADE	PART	ADE	PART
keelama	<i>forbid</i>	24	15	151	4
takistama	<i>hinder</i>	51	94	31	4
segama	<i>disturb</i>	10	17	13	
hoidma	<i>hold</i>		34		
hoiatama	<i>warn</i>		5		
pidurdama	<i>restrain</i>		1		
võõrutama	<i>alienate</i>		2		
aitama*	<i>help</i>			894	58
käskima*	<i>command</i>			280	31
laskma*	<i>allow</i>			2003	407
lubama*	<i>permit</i>			919	27
paluma*	<i>beg</i>			497	124
soovitama*	<i>suggest</i>			509	
võimaldama*	<i>enable</i>			584	7

TABLE 3. The distribution of infinitive constructions in the 1920s corpus

		MAST		DA	
		ADE	PART	ADE	PART
keelama	<i>forbid</i>		7	5	8
takistama	<i>hinder</i>		26	1	8
segama	<i>disturb</i>		1		
hoidma	<i>hold</i>		2		
hoiatama	<i>warn</i>				
pidurdama	<i>restrain</i>				
wõõrutama	<i>alienate</i>				
aitama	<i>help</i>			7	13
käskima	<i>command</i>			3	9
laskma	<i>allow</i>			11	193
lubama	<i>permit</i>			44	22
paluma	<i>beg</i>			2	308
soovitama	<i>suggest</i>			1	
wõimaldama	<i>enable</i>			7	1

Table 2 leads us to the following observations about the ICs in the 1990s:

Ade+DA (c. 5900) is dominant over Part+DA (c. 650).

Part+MAST (182) is more frequent than Ade+MAST (85).

From the data of the 1920s corpus in Table 3 we can make the following observations:

Part+DA (562) is dominant over Ade+DA (81).

Part+MAST (182) is the only construction attested, but Ade+MAST does not occur at all.

A neat and comprehensive explanation of these circumstances can be given if we assume that the change that took place in the ICs of Standard Estonian during the period under discussion is a composite process that can be reduced to three components:

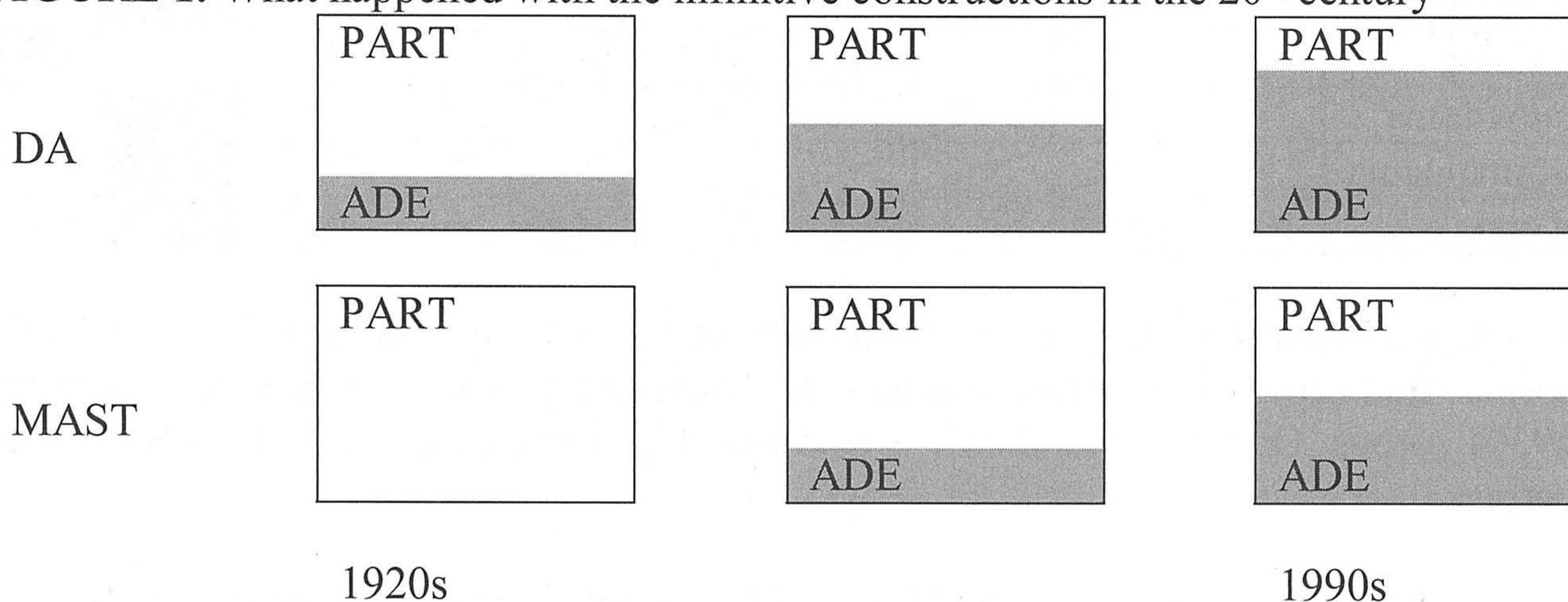
(I) While the use of the adessive (Ade+MAST / DA) has been on the rise, the use of the partitive (Part+MAST / DA) has been on a steady decline throughout the period.

(II) Within DA constructions, Part+DA used to be predominant in the 1920s, but the relation has been reversed to the advantage of Ade+DA during the course of the century.

(III) Within MAST constructions, Ade+MAST, which seems to have emerged in the 1920s at the earliest, has now become the equal of Part+MAST in verbs of 'forbidding.'

These three statements together comprise the hypothesis to be put to the validation test now. A graphical representation of the hypothesis may be given by Figure 1:

FIGURE 1. What happened with the infinitive constructions in the 20th century



5. In order to test the hypothesis, data are drawn from the five subcorpora of the Historical corpus. Tables 4 and 5 are chronologically arranged summaries of the historical corpus data.

For statistical purposes, one should admit that ICs are not sufficiently attested in some cells of the two tables. Nonetheless it is fairly clear that the historical corpus data do not exhibit any distribution patterns that defy the hypothesis under scrutiny.

TABLE 4. MAST constructions in 1890s–1950s

	1890s		1900s		1910s		1930s		1950s	
	ADE	PAR	ADE	PAR	ADE	PAR	ADE	PAR	ADE	PAR
keelama		9		1		8	1			1
takistama		1		1		6		5	2	3
hoidma		1		2		6		1		

Table 4 shows that Ade+MAST does not occur in the 1890s, 1900s, and 1910s corpora at all (cf. dark cells), whereas one occurrence of Ade+MAST is attested in the 1930s corpus and two in the 1950s corpus (cf. gray cells). This picture is consistent with the third statement of our hypothesis.

As for Table 5, the crossed cells indicate the absence of DA constructions, the light-gray background the dominance of Part+DA, and the dark-gray background the dominance of Ade+DA. The dominance of Part+DA over Ade+DA is evident in the corpora of the 1890s through the 1910s, whereas Ade+DA is apparently dominant in most verbs in the 1950s corpus. The 1930s data reveal a near balance of Ade+DA and Part+DA. Here the picture is nice and consistent with the first and second statements of our hypothesis.

TABLE 5. DA constructions in 1890s–1950s

	1890s		1900s		1910s		1930s		1950s	
	ADE	PAR	ADE	PAR	ADE	PAR	ADE	PAR	ADE	PAR
keelama	1	3		2	2	1	2	1	2	2
aitama	1	3	3	1	3	5	5	2	20	4
käskima	3	18	2	17	4	14	5	8	3	3
laskma	3	121	6	56	8	99	30	43	22	11
lubama	2	6	2	12	6	14	7	2	18	2
paluma		37		11		54		8	1	21
soovitama							5		12	
võimaldama							2		21	
Total	10	188	13	99	23	187	56	64	99	43

For the illustration, (2a) is one of our earliest examples of Ade+DA from the 1890s corpus, and (2b) is the earliest example of Ade+MAST in this study and is from the 1930s corpus. One may add that it is also the only occurrence of Ade+MAST in that corpus.

- (2) a. [...] seal oma jalga ja peaad **tohtril läbikatsuda** ja **rawitseda laske**. (1890s)
 ... get the doctor examine and cure your leg and head there.
- b. Mis te arvate, mis **keelaks** näiteks **minul** isiklikult seda **asutamast**? (1930s)
 What do you think would forbid me from starting it personally, for example?

FIGURE 2: The use of Part+DA and Ade+DA in the 20th century

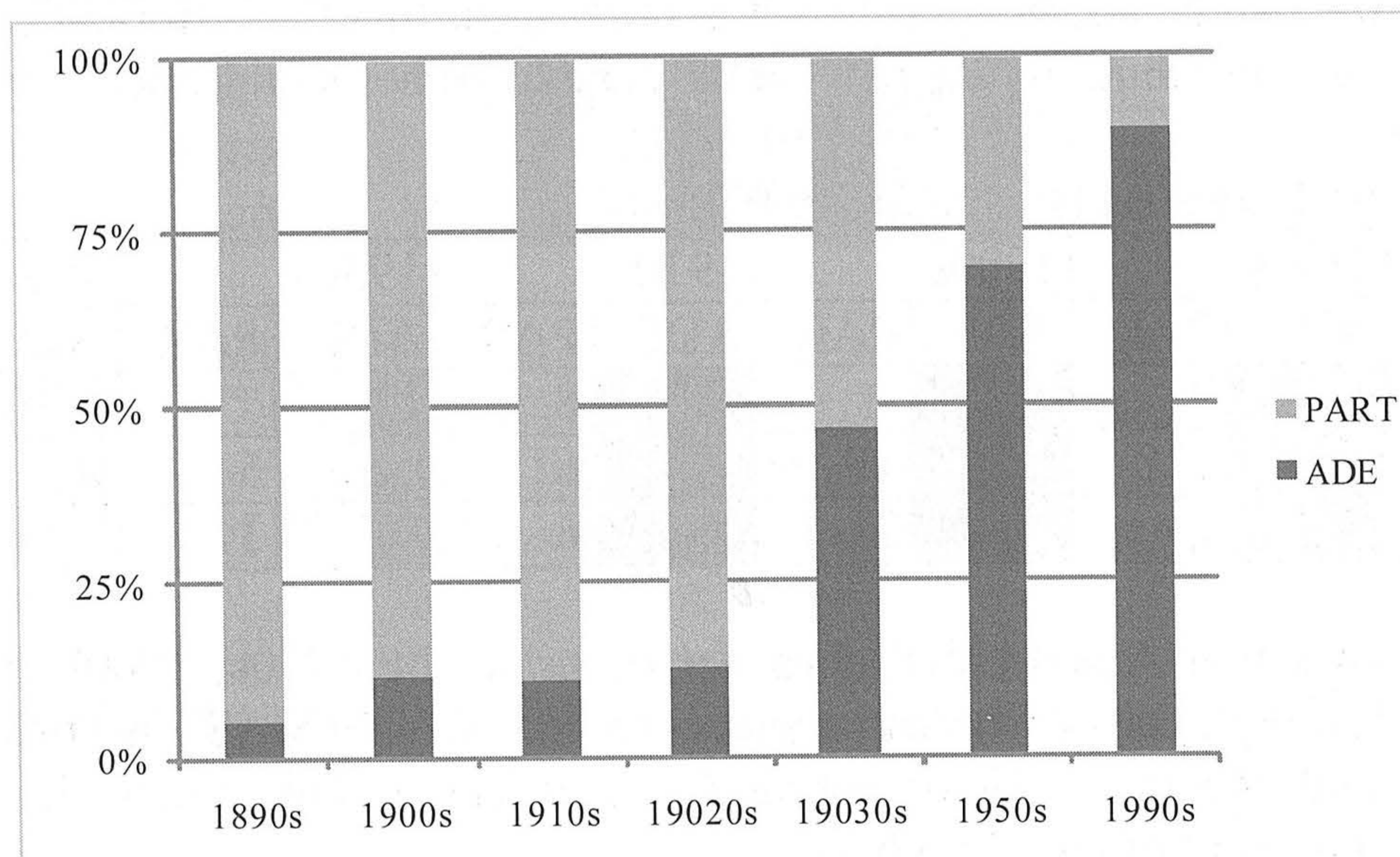


Figure 2 is a graphical representation of the change in the relative frequencies of Ade+DA and Part+DA. The steady advance of Ade+DA, or the gradual retreat of Part+DA for that matter, during the course of the 20th century is apparent.

6. We have yet to discuss the disturbingly aberrant behaviors of the two verbs in Table 5: *soovitama* 'to suggest' and *võimaldama* 'to enable.' There is one simple but perfectly plausible explanation for their anomaly. Namely, they were latecomers to the scene: it appears that *soovitama* and *võimaldama* started to occur in ICs in the early 1920s (cf. Table 3), when partitive constructions, Part+DA in particular, were losing their fertility and were already on the decline. The verbs took to Ade+DA from the very start because it was the only choice available to them if they were to join the community of the verbs of 'allowing' capable of taking ICs.

It is interesting to note that the verb *paluma* 'to beg' also exhibits an anomaly, though to a lesser degree and in a bit different way: it resisted the drift much longer than the other verbs and wasn't caught by Ade+DA until the 1950s.

This type of non-homogeneous spread of change in a language community reminds us of the phenomenon called *lexical diffusion*²⁴ of a phonetic change: a phonetic change does not usually affect all the relevant morphological items of the language simultaneously nor do all the members of the language community adopt the new pronunciation at the same time. As some words undergo

the change faster than others, and there may be quite a few words which have dual forms for some time. Though it is not clear whether the notion derived from the study of phonetic changes can be equally appropriately applied to grammatical changes as well, one may at least be allowed to speculate on the potential of language corpora as a tool of the study of grammatical changes.

²⁴See *inter alia*: Wang, William S-Y. 1969. Competing changes as a cause of residue. *Language* 45, 1969, pp. 9–25.

Redegunt

Sándor Csúcs
Nóra Falk
Viktória Tóth
Gábor Zaicz

Borító: Kemény Márton
Fotók: www.btk.ppke.hu

ISBN 978-963-88954-0-0 Ö
ISBN 978-963-88954-4-8 (Pars IV)

© Reguly Társaság, 2011

