

## Comparison of Two 12-Key Assignments on Mobile Devices

**V. Dagiene, G. Grigas, T. Jevsikova**

*Institute of Mathematics and Informatics, Vilnius University,  
Akademijos str. 4, LT-08663 Vilnius, Lithuania, e-mails: dagiene@ktl.mii.lt, grigas@ktl.mii.lt, tatjanaj@ktl.mii.lt*

### Introduction

Mobile devices, especially mobile phones, have become an attribute of almost everybody. They provide the new ways of accessing knowledge, co-operating with others and are used as a new media for open and distance learning [1, 2]. The quality of distant learning depends not only on the quality of learning materials and methods, but as well on tools and communication used, e.g. mobile devices [3]. SMS messaging is one of the most popular mobile services. According to worldwide statistics, over a billion of SMS messages are sent daily. So, it is crucial that the keypad, as one of the main components of human-computer interface of mobile devices, is natural and friendly.

In 2003, ETSI, the European Telecommunications Standards Institute, published the first version of standard [4] for letters, digits and special characters, covering the official languages of the European Union (EU), European Free Trade Association (EFTA) members, Russia, as well as countries with applicant status for EU at that time. The main issues of the standard was repertoire of characters, including letters of alphabet for languages, covered by the standard, their sorting and layout on the 12-key keypad of mobile devices. In 2007, the second version of the standard mentioned above was developed [5]. Its rationale was described in the papers [6; 7]. The standard was considerably extended to cover other major languages, spoken in Europe, including the official languages, minority languages and immigrants' languages. The principles of character layout remain the same as in the first version of the standard.

Before that and till now, mobile phones have also been produced with a slightly different character layout for languages, based on Latin script. So, it is reasonable to compare these two layouts.

This research has been partly carried under the framework of the project "Science for Business and Society" (Project number: VP2-1.4-UM-03-K-01-019).

### Two variants of key assignment

There are two widespread variants of key

assignments on 12-key keypad of mobile devices for languages using Latin script. They differ in sequence of characters, obtained by multiple pressing of the same key with short time interval. The characters appear in the following sequence (Variant 1 and Variant 2).

*Variant 1:*

- Basic Latin letters;
- Digit;
- Other Latin letters from the alphabet of a particular native language and letters of some foreign languages.

*Variant 2:*

- Alphabet letters of particular native language;
- Digit;
- Letters of some foreign languages.

The first variant is defined by the ETSI standards; the second variant is used in mobile phones, produced by Nokia. As an example, let us give the assignments for three languages, two Baltic and one Slavic (Table 1):

1. Lithuanian, with moderate number of letters, distinguished from the basic Latin alphabet (letters QWX are not used, and 9 extra letters are used: AČĖĖIŠŪŪŽ);

2. Latvian, with greater differences from the basic Latin alphabet (QWXY are not used, and 13 extra letters are used: ĀČĒĢĪĶĻŅŖŠŪŽ);

3. Polish (V is not used, and 9 extra letters are used: AĆĘŁŃÓŚŻ).

### Properties of two variants

Let us examine the differences of the two variants of key assignment in more detail according three properties:

1. *Digit inside the letter sequence.* According to the variant 1, digits are included inside the sequences of native letters on some keys (on 6 keys (2, 3, 4, 7, 8, and 9) for Lithuanian, on 8 keys for Latvian, and on 6 keys (2, 3, 5, 6, 7, and 9) for Polish). According to the variant 2, all the digits go after letters of native alphabet.

A digit inside the letter sequence may be considered as a separator. In the variant 1 it splits the sequence of native letters, assigned to one key, into two parts: those are inside the basic Latin and those are outside of the basic

Latin, usually letters with diacritics. In the variant 2, the digit separates native letters from the foreign letters.

Splitting sequence of native letters into two parts by form (without diacritics and with diacritics) is unnatural

from the user's point of view. Splitting all letters into native and foreign is natural and desirable. This is in favour of the variant 2.

**Table 1.** Key assignments for Lithuanian, Latvian and Polish languages

Key	Variant 1			Variant 2		
	Lithuanian	Latvian	Polish	Lithuanian	Latvian	Polish
2	ABC2AČ	ABC2ĀČ	ABC2AĆ	AABCČ2ĀĀĀÆ	AĀBCČ2ĀĀĀÆ	AABCČ2...
3	DEF3EĖ	DEF3Ē	DEF3Ę	DEĖĖF3ĒĒ	DEĒĒF3ĒĒĒ	DEĖF3...
4	GHI4Į	GHI4ĢĪ	GHI4	GHI4ĢĪ	GĢHIĪ4Į	GHI4...
5	JKL5	JKL5ĶĻ	JKL5Ł	JKL5ĶĻĒ	JKĶĻĻ5Ē	JKL5...
6	MNO6	MNO6Ņ	MNO6ŃÓ	MNO6ŅŃŃŃ	MNOŅ6ŃŃŃ	MNŃÓ6...
7	PQRS7Ś	PQRS7RŠ	PQRS7Ś	PQRSŚ7RBS	PQRŔŚ7BS	PQRSŚ7...
8	TUV8Ū	TUV8Ū	TUV8	TUŪV8Ū	TUŪV8Ū	TUV8...
9	WXYZ9Ž	WXYZ9Ž	WXYZ9ŻŻ	WXYZŽ9	WXYZŽ9	WXYZŽ9...

2. *Order of letters.* Let us show natural alphabet order, presented in ETSI standards, and ordering imposed by key assignments in both variants.

*Lithuanian* (with foreign letters Q, W, Z included):

natural: AABCČ DEĖĖF GHIĮY JKL MNO PQRSŚ TUŪV WXYZŽ

variant 1: ABCAČ DEFĖĖ GHIĮ JKL MNO PQRSŚ TUVŪ WXYZŽ

variant 2: AABCČ DEĖĖF GHIĮ JKL MNO PQRSŚ TUŪV WXYZŽ

*Latvian* (with foreign letters Q, W, X, Y included):

natural: AĀBCČ DEĒĒF GĢHIĪ JKĶĻĻ MNŅO PQRŔŚ TUŪV WXYZŽ

variant 1: ABCĀČ DEFĒ GĢHIĪ JKĶĻĻ MNŅO PQRŔŚ TUVŪ WXYZŽ

variant 2: AĀBCČ DEĒĒF GĢHIĪ JKĶĻĻ MNŅO PQRŔŚ TUŪV WXYZŽ

*Polish* (with foreign letter V included):

natural: AABCČ DEĖĖF GHI JKL MNŃÓ PQRŚ TUV WXYZŻ

variant 1: ABCAČ DEFĖ GHI JKL MNŃÓ PQRŚ TUV WXYZŻ

variant 2: AABCČ DEĖĖF GHI JKL MNŃÓ PQRŚ TUV WXYZŻ

*Note.* Lithuanian letter Y is not shown as deviated because the reason for deviation is another, and not common to other languages.

Letters, which are deviated from natural order, are shown in bold typeface. Deleting them from the sequence will restore the alphabetic order of remaining letters.

In the variant 2, letters on all keys are ordered alphabetically.

In the variant 1, for all three languages there is some deviation from the alphabetic order as shown in Table 2.

All this shows that this property is in favour of variant 2.

Let us note, that there are some proposals [8] to deviate from the alphabetic order on the same key (e.g. reorder ABC to ACB, because the frequency of the letter C is higher than that of B) in order to increase the typing speed. However, in our case deviation from alphabetic order decreases the typing speed, because skipping through

digit requires extra key pressing. This is in favour of the variant 2

**Table 2.** Deviation from the alphabetic order in variant 1

Parameter	Lithuanian	Latvian	Polish
Number of keys with destroyed alphabetic order	3	7	3
Total number of letters not in alphabetic order	5	7	3

3. *Uniformity of language-independent sequences of letters.* In the variant 1, all sequence of basic Latin letters (e. g. ABC, DEF) is the same on any mobile phone, independently of the language.

In the variant 2, such common sequences are found only for some keys (e. g. 5 and 6 for Lithuanian).

*Producer vs. user.* Let us consider the three properties, discussed above (digit inside the letter sequence, order of letters, and uniformity of language-independent sequences of letters) from the software/hardware manufacturer's point of view, as well as from the user's point of view. Software/hardware manufacturer will obviously be interested in property 3 (the same production template can be used for many languages), and so will support variant 1. The property 3 may be considered useful only for devices, shared by users who speak different languages. But it is not the case for mobile phones as individual device in its nature. Properties 1 and 2 are both important from the ordinary user's viewpoint, and are in favour for variant 2. If we support an idea that the producer works for user, then it would be reasonable to accept variant 2.

*Cultural issues.* Mobile phones play an important role in our everyday life. Amount of texts, sent via SMS, is so large that they have considerable influence on our language habits and writing culture. Moving some letters to the secondary, less convenient position imposes an idea that these letters are less important and may be ignored, especially when typing in a hurry or in not convenient conditions, e. g. in a street, on a bus, etc. Similar situation on ignorance of letters with diacritics was observed when



Comparing Table 1 and Table 3, we can see that the number of foreign letters in the variant 2 is greater than this in variant 1. Both layout methods do not restrict the set of letters directly. However, the fact that there are no separator between native letters with diacritics and foreign letters may have some indirect influence to restrict the number of foreign letters. ETSI standard does not present any foreign letters after digits for Lithuanian, Latvian, Polish and some other languages.

According to the standard letters of Cyrillic, Greek, and other scripts, except Latin, and digits on the keys must be placed in such order: set A, digit, set B, i. e. using the second assignment variant.

According to the standard, letters of Latin script and digits on the keys must be placed in such order: basic Latin letters, digit, other letters of a particular language together with letters of some foreign languages, i. e. using the first assignment variant.

Thus the division of letters into two sets (A and B) is ignored, and so we see the contradiction inside the standard itself.

Moreover, the standard insists inequality of languages using Latin script (excluding English) against the languages using other scripts.

In the number of languages (14 of 25 or 56% from Table 3), some basic Latin letters are not used as native letters: (Q and W in 10, X and Y in 7, C and K in 2 languages, V in 1 language). Commonly they are used only in foreign names. According to the ETSI standard, they belong to group B for those languages. However, the standard requires that all basic Latin letters, regardless whether they are present in the alphabet of particular language or not, must be allocated before digits, i. e. in the same position as native letters from group A. This is another contradiction in the ETSI standard itself for the majority of languages using Latin script.

### The role of letters outside the basic Latin subset

The role of letters outside the basic Latin subset may be evaluated by the number of such letters in the alphabet and their frequency in texts. The number of such letters may be calculated from Table 3. The frequencies of all such letters in three languages, considered in Table 1, are as follows: Latvian 12% (calculated ad hoc), Lithuanian 8,5% [11], Polish 7% [12].

Ranking of letters, according to frequencies of letters in those three languages, is as follows:

lv: ISETURN ĀKDMVLJPOZĒĪBGŠCŅŪĻŽĢĢĶRŌH,

lt: IASEOTRNUKMLPVDVJGB ŪYŠŽĄĲĖĈĈZĖFŪH,

pl: AIOEZNRWSTCYKDPMUJL ŁBGĘHAÓŹŚĆFNQZVX.

Blank space in the sequences denotes the place before first non-basic Latin letter.

There are no reliable statistics on letter frequencies of all languages, listed in Table 3. However, ranking of letters may be found at [13]. So, we use it here.

Some data, related to the role of letters outside the basic Latin for a number of languages, are presented in Table 4.

The percentage of non-basic Latin letters (column b) against the number of all letters (column a) in the alphabet is presented in the column c.

In the column f, the percentage of the number of basic Latin letters (column e) preceding the first non-basic Latin letter (column d) in the ranked sequence of all letters (column a).

Note that the last Fig. shows the role of non-basic Latin letters in opposite direction: the lower number, the larger role.

**Table 4.** Usage of letters outside the basic Latin subset

Lng. code	Language name	Number			Ranking		
		All	N	%	1 <sup>st</sup>	N	%
		a	b	c	d	e	f
Cs	Czech	38	15	39	Í	7	18
Da	Danish	29	3	10	Ø	19	66
De	German	30	4	13	Ü	21	70
Es	Spanish	31	7	23	Í	17	55
Et	Estonian	27	6	22	Õ	17	63
Fi	Finish	28	2	7	Ä	7	25
Fr	French	39	13	33	É	15	38
Ga	Irish	42	16	38	Í	12	29
Hu	Hungarian	34	9	26	Á	13	38
Is	Icelandic	32	9	28	Ð	8	25
It	Italian	31	7	23	Ò	21	68
Lv	Latvian	35	13	37	Ā	8	23
Lt	Lithuanian	32	9	28	Ū	18	56
Mt	Maltese	28	4	14	Ħ	19	68
Nl	Dutch	31	6	19	Ë	23	74
No	Norwegian	30	4	13	Ø	19	63
Pl	Polish	32	9	28	Ł	19	59
Pt	Portuguese	36	13	36	Ã	20	63
Sv	Swedish	30	4	13	Ä	14	47
Tr	Turkish	30	7	23	İ	6	20
Average		32	8	24		15	49
Max		42	16	39		23	74
Min		27	2	7		6	18

We can see that the average number of basic letters (group A, according to ETSI standard) with diacritics and other letters, not present in basic Latin, is 8 in 20 European languages using Latin alphabet and varies from 2 (Finish) to 16 (Irish). The first letters in the ranked frequency sequences of all languages belong to the Latin basic subset. However, the first letter from outside of basic Latin (dotless I, Turkish alphabet) appears in the 7th position. In average, these letters appear in the middle (after 49%) of the ranked sequences of the alphabets.

The Fig.s, presented above, show that the role of non-basic Latin letters is important, and they must be treated truly as basic letters without any restrictions comparing with other letters.

### Labels on keys

ETSI standards do not define key labelling. One of good principles of interface design is WYSIWYG (What

You See Is What You Get). Proper labelling excludes uncertainty. Uncertainty avoidance is the first principle of cultural dimensions in Hofstede's and others intercultural management model [14; 15].

So, it would be natural to have engraved key labels, such as AĄBCCĆ, DEĘĖF... for Lithuanian, AĀBCCĀ, DEĒĒF for Latvian, AÄBC for German, etc. (Table 5). Of course, it is yet another concern for mobile phone producers, but it makes a device more user-friendly.

**Table 5.** Letter labels on keys

Digit	Lithuanian	Latvian	Polish
2	AĄBCCĆ	AĀBCCĀ	AĀBCCĀ
3	DEĘĖF	DEĒĒF	DEĒĒF
4	GHIĮ	GĢHIĪ	GHI
5	JKL	JKĶLLĻ	JKLL
6	MNO	MNŅŅO	MNŃŃOÓ
7	PQRSŠ	PQRŅŠŠ	PQRSSŚ
8	TUŲŪV	TUŪV	TUŪ
9	WXYŽŹ	WXYŽŽ	WXYZZŹ

We can see that the number of letters on one key grows up to 6 on 7th (for Latvian) and 9th (for Polish) key. It is natural: these languages have more letters. Average number of native letters is 32 and maximum number is 42 (Table 4). Thus 6 letters on one key are not so many. In extreme cases, some letters can be moved to a second row (as for example Cyrillic or Greek letters) or can be replaced by some thin surrogate symbol.

## Conclusions

Analysis of two variants of key assignments has shown that letter layout for languages, using Latin script on 12-key keypad of mobile devices, where all letters of a particular language are placed before digits, ensures more natural interface between the device and human being than the layout where letters with diacritics are separated from their basic counterparts and moved after digits.

The principle of WYSIWYG may be ensured, placing key labels with all letters, assigned to that key.

There are contradictions in ETSI standard itself, caused by not equal treatment of letters with diacritics and other native letters, and treatment of all letters of basic Latin alphabet as native letters.

Implementation of statements, expressed in paragraphs above, would make the rules of layout more uniform for the languages of various scripts.

Equal treatment of all letters of the alphabet ensures consistency with user's cultural and language environment.

## References

1. **Cornu B.** New Media and Open and Distance Learning: New Challenges for Education in a Knowledge Society. // *Informatics in Education*, 2007. – Vol. 6. – No. 1. – P. 43–52.
2. **Hölbl M., Welzer T.** Students' Feedback and Communication Habits using Moodle. // *Electronics and Electrical Engineering*, 2010. – No 6(102). – P. 63–66.
3. **Targamadžė A., Petrauskienė R., Rublauskas D.** Influence of Technologies on Quality of Distance Learning. // *Electronics and Electrical Engineering*, 2010. – No 6(102). – P. 131–134.
4. **ETSI ES 202 130 v.1.1.1.** Character repertoires, orderings and assignments to the 12–key telephone keypad (for European languages and other languages used in Europe). – ETSI, 2003.
5. **ES 202 130 v.2.1.2.** Character repertoires, orderings and assignments to the 12–key telephone keypad (for European languages and other languages used in Europe). – ETSI, 2007.
6. **Böcker M., Larsson K. L., von Niman B.** Increasing the Usability of Text Entry in Mobile Devices for European Languages and Languages Used in Europe // Aykin, N. (Ed.) *Usability and Internationalization. Global and Local User Interfaces*, pt 2, proc. Lecture Notes in Computer Science. – Vol. 4560. – Springer Berlin / Heidelberg, 2007. – P. 13–21.
7. **Böcker M., Larsson K. L., von Niman B.** Standardization Supporting Cultural Diversity: Character Repertoires, Ordering and Assignment to the 12–key Telephone Keypad for European Languages and Languages Used in Europe // *Proceedings of the 8th conference on Human–computer interaction with mobile devices and services, ACM International Conference Proceeding Series*, 2006. – Vol. 159. – P. 281–282.
8. **Ha R. W., Pin-Han Ho, Huemin Shen.** SIMKEYS: An Efficient Keypad Configuration for Mobile Communications. // *IEEE Communications Magazine*, 2004. – Vol. 42. – Iss. 11. – P. 136–142.
9. **Grigas G., Pedzevičienė S.** Klaviatūros įtaka rašybos klaidoms, susijusioms su rašto ženklų vartojimu [The Relation Between Keyboard Layout and Spelling Errors of Particular Language Specific Letters] // *Informacijos mokslai*, 2009. – T. 50. – P. 200–204.
10. **Schäler R.** The cultural dimensions in software localisation. // *Localisation Focus*. – Vol. 1. – Iss. 2. – Localisation Research Centre, Limerick, 2002.
11. **Žilinskienė V.** Lietuvių kalbos dažninis žodynas [Frequency vocabulary of Lithuanian language], 1990. – Vilnius: Mokslas.
12. **Przepiórkowski A.** Poradnia językowa [Language data], 2006. Online: <http://poradnia.pwn.pl/lista.php?id=7072>.
13. **Kelk B.** Letter frequencies (ranking for various languages), 2007. Online: <http://www.bckelk.ukfsn.org/words/etaoin.html>.
14. **Hofstede G.** *Culture's consequences*. – London: Sage Publications Ltd. – 1980. – 13 p.
15. **Caganova D., Cambal M., Weidlichova Luptakova S.** Intercultural management – trend of contemporary globalized world. // *Electronics and Electrical Engineering*, 2010. – No 6(102). – P. 51–54.

Received 2010 10 13

**V. Dagiene, G. Grigas, T. Jevsikova. Comparison of Two 12-Key Assignments on Mobile Devices // Electronics and Electrical Engineering. – Kaunas: Technologija, 2011. – No. 2(108). – P. 97–102.**

Two variants of key assignments on 12-key mobile devices for languages based on Latin script are discussed. It is shown that the variant where all letters of the alphabet of a particular language are placed in front of digits is more preferable than that where letters with diacritics are placed after digits. The investigation was conducted, using Lithuanian, Latvian and Polish alphabets as the examples.

However, all the results and conclusions, presented in the paper, are valid for any Latin script language, using the alphabet with more than 26 letters of basic Latin set. Ill. 1, bibl. 15, tabl. 5 (in English; abstracts in English and Lithuanian).

**V. Dagienė, G. Grigas, T. Jevsikova. Simbolių išdėstymo dvylikos klavišų mobiliuosiuose įrenginiuose variantų palyginimas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2011. – Nr. 2(108). – P. 97–102.**

Straipsnyje nagrinėjami du ženklų išdėstymo dvylikos klavišų mobiliuosiuose įrenginiuose variantai. Įrenginiai skirti lotynų abėcėlę vartojančioms kalboms. Parodyta, kad variantas, kai tam tikros kalbos abėcėlės visos raidės pateikiamos prieš skaitmenis, yra priimtinesnis negu variantas, kai raidės su diakritiniais ženklais yra pateikiamos po skaitmenų. Tyrimas atliktas remiantis lietuvių, latvių ir lenkų abėcėlėmis, tačiau visi straipsnyje pateikti rezultatai ir išvados galioja ir visoms kitoms lotynų abėcėlę vartojančioms kalboms, turinčioms daugiau kaip 26 pagrindinės lotynų abėcėlės raides. Il. 1, bibl. 15, lent. 5 (anglų kalba; santraukos anglų ir lietuvių k.).