

# A Myriad of Ways to Say: “Wear a mask!”

**Cvetana Krstev**

University of Belgrade  
Faculty of Philology  
cvetana@matf.bg.ac.rs

**Duško Vitas**

University of Belgrade  
Faculty of Mathematics  
vitas@matf.bg.ac.rs

## Abstract

This paper presents a small corpus of notices displayed at entrances of various Belgrade public premises asking those who enter to wear a mask. We analyze the various aspects of these notices: their physical appearance, script, lexica, syntax and style. A special attention is paid to various obligatory and optional parts of these notices. Obligatory parts deal with wearing masks, keeping the distance, limiting the number of persons on premises and using disinfection. We developed local grammars for modelling phrases that require wearing masks, that can be used both for recognition and for generation of paraphrases.

**Keywords:** short messages, local grammars, phrase generation, paraphrasing.

## 1 Introduction

Short messages have been attracting attention of linguists and researchers in natural language processing (NLP) for some time. One of the reasons is that it is a widespread type of communication, and is not limited to use among young people and entertainment. It has been noticed that short messages use a specific language and a particular style. For that reason, numerous corpora of short messages have been collected and can be explored by means of NLP tools. A corpus of 88,000 French SMS messages was collected, anonymized and made available for research purposes (Panckhurst, 2017). Petrović et al. (2010) presented a large Twitter corpus of 97 million posts and made it publicly available for researchers working in social media, NLP and large-scale data processing.<sup>1</sup>

In many cases researchers tailored their own corpora of short messages suiting their purposes. Bernicot et al. (2012) compiled a corpus of 864

<sup>1</sup>The corpus is no longer available due to change in Twitter policy.

SMS messages produced by French-speaking adolescents in order to analyze the effects of writers' characteristics on message length, dialogue structure, and message function. A corpus of French Twitter posts containing complaints regarding railway services was used to investigate linguistic directness and indirectness and differentiate them from perceived (im)politeness (Depraetere et al., 2021).

Graffiti on the walls of urban spaces are also a type of short messages; moreover, they have a much longer history than messages on today's social media. They have been analyzed from various perspectives: political, sociocultural, and linguistic (Alonso, 1998). The graffiti for analysis are often selected from a specific area, such as a university campus. Authors applied corpus method to analyse 378 graffiti found on walls of two Jordan universities (Al-Khawaldeh et al., 2017), and found that they express different themes: personal, social, national, religious, political etc.

A formulaic short messages are often sent on the occasion of Christmas and New Year. Christmas and New Year messages issued by important persons, like heads of state, which are far from being formulaic, have received more than their fair share of attention (Sauer, 2007). Still, some researchers were interested in the formulaic expressions: Deng et al. (2010) analysed how the language of Chinese SMS messages conveying Christmas wishes reflects a shift in cultural values and customs, while Włosowicz (2011) analysed how foreign language learners' mother tongue and cultural background influence their Birthday, Christmas and New Year's wishes.

Linguists showed interest in short messages written in Serbian as well. The use of shortening, clipping and elliptic constructions in text messages was analysed by several authors (Polovina and Jelić, 2020; Jelić and Vekarić, 2019). Graffiti

that emerged in Serbia in its transition era were analysed and messages were classified in overarching categories, as hate messages and love messages (Mršević, 2014). NLP specialists collected corpora of various types of short messages in order to solve different problems. Šandrih (2018) prepared a corpus of 5,500 Serbian SMS messages to test the system for detecting their sender. A corpus of 9,059 Serbian Twitter posts was collected in order to determine how their sentiment is affected by the use of negation (Ljajić and Marovac, 2019). Jokić et al. (2021) collected and manually annotated for hate-speech 6,436 tweets to be used for training hate-speech detection applications.

In this paper we are interested in notices that announce precaution measures related to the Covid-19 pandemic at front doors of public premises. These notices are similar to graffiti as they are public and are not a product of social media. However, contrary to graffiti, their content is restricted and in that respect they are closer to Christmas and New Year wishes.

Our paper is organized as follows. Section 2 presents a small corpus of notices related to precaution measures against Covid-19 virus. In Section 3 the lexica used in these notices is analysed. Sections that follow are restricted to the part of these notices that concern protective masks: their basic syntactic patterns (Section 4) and their semantic value (while in Section 5). Finally in Section 6 we show by generating mask messages of determined structures that there are myriad ways to say: “Wear a mask!”. In Section 7 we discuss avenues for future research.

## 2 About the Corpus

Our corpus is based on notices that were photographed between 21 January 2021 and 30 March 2022 and then re-typed. We considered as a single notice everything written on one sheet of paper. Not all notices were originally composed; instead, many were acquired from Internet and other sources, and used on entrances to many different facilities. We filtered only different notices from all notices photographed. We do not claim that our set of notices is in any way representative since all notices were photographed in the center of Belgrade at the walking distance from the place of the residence of the authors of this paper. The total number of photographed notices is 231.

**Physical appearance** – The majority of notices

were printed on a paper (207), 23 notices were handwritten. There was one 3D notice – a word “Obavezno” (obligatory) written on an actual mask.

**Capitalization** – The majority of notices (165) were written using only upper-case letters, 45 notices were written using lower-case letters, while in 21 notices only some parts were written using only upper-case letters for emphasis.

**Script** – The majority of notices (187) were written using Latin script. Among them 18 did not use diacritics. 44 (19%) of notices were written in Cyrillic script. Among 23 handwritten notices, 8 (34.7%) were written in Cyrillic. These findings are in line with the conclusions reached by Ivković (2013) that the Latin alphabet dominates over Cyrillic on Serbian news websites and the use of non-standard Latin orthographic variants (diacritics omission) is becoming stable.

**Emphasizing** – Various methods of emphasis were used. Many notices included images (of a mask, people maintaining a distance, etc.) – 68 (29.4%) such notices were in the selected set. Other means of emphasizing were: using bold font (18), underlining (23), colors (mostly red) (29), increasing the font size (13). In many cases more than one mean of emphasizing was used. The other means of emphasizing the message of the notice was the use of the exclamation mark. It was used in 52 notices, either only once (40) or repeatedly (two times, three times, and up to 16 times). It is interesting to note that only once an emoticon was used (a smiley).

**Multilinguality** – All notices were written in Serbian. However, some of them had translations in English, and they vary in form as much as those written in Serbian that we will explore in following sections. These 5 notices in English concerning the use of masks were: “*Please wear protective mask in public areas*”, “*No entry without face mask*”, “*use a protective mask*”, “*Face mask required*”, “*Please don’t enter without a face mask*”.

In some notices images were intertwined with words to convey the meaning. For instance, <img of a person> 5 <img of a mask> Hvala (Thank you). Such notices were excluded from further analysis, as well as notices or their parts written in English.

The length of notices is between one word (a word “Obavezno” (obligatory) written on an actual mask) and 84 words. The average length of notices is 14.3 words, while two thirds of them were

written using no more than 15 words.

Each notice mentioned one to four protective measures: (a) wearing a mask – 207 notices; (b) number of persons allowed on premises – 64 notices; (c) keeping the distance – 39 notices; (d) disinfection – 17 notices.

The majority of notices (157) listed only one protective measure, 49 notices listed two, 17 listed three, while 8 notices listed all four measures. The parts of notices concerning masks were the shortest (6.74 words), followed by disinfection (8.29 words), number of people on premises (8.39 words), and keeping a distance (8.74 words).

Besides parts of notices listing the protective measures, some of them have one or more additional parts. They are:

**Attracting attention and addressing those who enter** – This part was represented in 70 messages (30.3%). It is simple and did not vary much in form. To attract attention several words were used: *Obaveštenje* (Notice) (12), *Pažnja* (Attention) (9), *Važno* (Important) (2), *Stop* (1). Three forms were used to address those who enter premises: *Poštovani*, (Respected,) (11) – this is a very formal and impersonal form of address that is often used for written official communication with unknown persons; A slightly less formal *Poštovani kupci*, (Respected customers,) (30) where the most frequently used word *kupci* can be replaced with *potrošači* (consumers), *klijenti* (clients), *posetioci* (visitors), *gosti* (guests), *sladokusci* (gourmand); *Dragi kupci*, (Dear customers,) (5) where the word *kupci* can be replaced with some of the previously listed words – the more informal form of address, but still very polite. These findings lead us to the question of whether Serbian society still belongs to “solidarity cultures” of the East rather than “distance cultures” of the West (Schlund, 2014). In some notices parts for both attracting attention and addressing customers were used.

**Invoking authority and explanations** – This part occurred in 47 of selected notices (20%). It is rather long and without a strict form. It conveyed the reasons for the protective measures and/or who has prescribed them. The statement expressing reasons for prescribing necessary measures started usually with *zbog/usled pandemije...* (due to the pandemic...) or *u cilju/radi sprečavanja pandemije* (in order to prevent pandemics...). The statements invoking authorities started with *Po/Prema/Na osnovu/U skladu sa odlukom...* (According/On the

basis/In accordance with the decision...). The most frequently mentioned authority is the government of the Republic of Serbia (14), and besides it the Ministry of Health (1) and *Krizni štab* (Crisis Response Team) (1). In the cases when a specific authority is not mentioned, a particular decision or regulation published in *Službeni glasnik* (Official Gazette) is listed (3). In one case the precise article of the regulation is mentioned with no less than 13 issues of the Official Gazette. Both messages – authority and explanation – sometimes occur together: *Usled odluke Vlade Republike Srbije, a u cilju suzbijanja epidemije...* (Due to the decision of the Government of the Republic of Serbia, and in order to suppress the epidemic...). It is legitimate to ask which of the features mentioned by Njegovan et al. (2011): credibility, exclusiveness, uniqueness, omnipresence, validity that characterize bureaucratic authority, are most likely to appeal to customers to obey to precaution measures.

**Gratitude** – This part occurred in 51 of selected notices (21.6%). It is simple and does not vary much in form. The used expressions are: *Hvala* (Thanks) (23) – neutral, *Hvala Vam* (Thank you) (1) – a slightly more personal, *Hvala (Vam) (lepo)/Zahvaljujemo se na razumevanju* (Thanks/We thank (you) (nicely) for your understanding) (22) – apologetic (we apologize that you have to wear a mask and we thank you for understand it), *Hvala unapred/Unapred hvala* (Thanks in advance) (3) – appealing to customers’ conscience because they accept the gratitude before they have done what is asked of them. Finally, the form *Zahvaljujemo što poštujete navedenu meru* (We thank you for complying with this measure) was used twice. It is interesting to note that the form *Hvala lepo* (Thank (you) nicely) once popular in everyday communication was encountered only once.<sup>2</sup>

**Miscellaneous** – Occasionally some miscellaneous information was added to notices (38 cases). This information was sometimes completely unrelated to the precaution measures, e.g. working hours of a shop. In a number of cases some additional information is added to a certain precaution measure, like *Zadržavanje u radnji je do 10 minuta!* (Staying in the shop is up to 10 minutes!) or *Ukoliko nemate masku, dobićete je u knjižari...* (If you don’t have a mask, you will get it in the bookstore...). In some notices (24) additional statements

<sup>2</sup>This form has 54 occurrences in the SrpKor2013 (<http://www.korpus.matf.bg.ac.rs/>), and 4,622 in the SrpKor2021 (<https://noske.jerteh.rs>).

NOUN	Freq	VERB	Freq	ADJ	Freq	ADV	Freq
<i>maska</i>	207	<i>moliti</i>	48	<i>obavezan</i>	122	<i>obavezno</i>	25
mask		to request		mandatory		mandatorily	
<i>nošenje</i>	91	<i>dozvoliti</i>	27	<i>zaštitni</i>	84	<i>istovremeno</i>	16
wearing		to permit		protective		simultaneously	
<i>hvala</i>	50	<i>moći</i>	27	<i>poštovan</i>	40	<i>maksimalno</i>	7
gratitude		can		respected		maximally	
<i>ulazak</i>	49	<i>nositi</i>	25	<i>drugi</i>	16	<i>najmanje</i>	7
entering		to wear		other		at least	
<i>objekt</i>	39	<i>držati</i>	19	<i>maksimalan</i>	11	<i>najviše</i>	7
facility		to keep		maximal		at the most	
	<b>1272</b>		<b>291</b>		<b>417</b>		<b>82</b>

Table 1: The most frequent nouns, verbs, adjectives and adverbs in the Mask corpus

were used as encouragement for people to respect imposed measures: *Budimo odgovorni* (Let's be responsible) or *Čuvajmo sebe i druge* (Let's take care of ourselves and others).

**Signature** – 42 notices were signed by the facility which attached a notice. Even this part of selected notices was not completely uninteresting. More than half of the signed notices (22) reveal appealing foreign or foreign-like firm names: *Beauty and the beast center*, *Beomelody d.o.o.*, *Ušće Shopping center*.

Serbian has a binary pronominal system of address which employs one pronoun (second person singular – *Ti*) for familiar address and another (second person plural – *Vi*) for formal address. Although studies (Milosavljević, 2018) have shown that the informal address has been gaining attraction in everyday communication over the last couple of decades, e.g. in media were a show host addresses a guest, in our notices a familiar address using the *Ti* pronoun was never used. Some statements are ambiguous, such as *Molimo vas da se pridržavate mera zaštite radi sprečavanja širenja zaraze koronavirusom* (Please adhere to protection measures to prevent the spread of coronavirus infection), which can refer both to one person addressed by the *Vi* pronoun and to a group of people. However, a statement like *Obavezno nosi<sub>sing</sub> masku<sub>sing</sub>* (Be sure to wear a mask) was not found in our selection of notices.

### 3 Lexica used in notices

When preparing our tiny little corpus the notices were typed as they were and corrected only evident typos, leaving grammatical and orthographic errors. The corpus consists of 3,581 tokens and

3,285 words. Among words, there were 193 different nouns, 49 different verbs, 57 different adjectives and 15 different adverbs. The five most frequently used nouns, verbs, adjectives and nouns are listed in Table 1.

In our sample specific groups of nouns were identified. The first group containing 34 nouns were used to refer to places to which notices about protective measures apply. Here one can distinguish the most general concepts: *objekat* (facility) (39), *prostor* (space, area) (13), *prostorija* (room) (9), *zgrada* (building) (2), *mesto* (place) (1). These concepts can be further qualified: *prodajni prostor* (shopping area), *poslovni prostor* (business area), *javni prostor* (public area), *zatvoreni prostor* (enclosed area) *prodajni objekat* (shopping facility), *maloprodajni objekat* (retail shopping facility), *radna prostorija* (working space), *javno mesto* (public place). The more specific concept was represented by three nearly synonymous words *radnja* (32) *prodavnica* (12) (shop), *maloprodaja* (retail) (1). The remaining 26 nouns were used to name a place for a specific activity, like *menjačnica* (exchange office), *pekara* (bakery), *fakultet* (faculty).

The other group of specific nouns refers to people to whom notices are addressed. Here also very general concepts were used, lexicalized by *osoba* (39) and *lice* (8) (person), *gradjanin* (1) (citizen), *ljudi* (people) (4). Besides them, more specific concepts were used for potential: *kupac* (shopper) (28), *potrošač* (consumer) (14), *mušterija* (customer) (3), while other concepts were specific to particular activity, like *gost* (guest) (1) and *posetilac* (visitor) (7).

It should be noted that concepts referred to by these two specific groups of nouns are related in



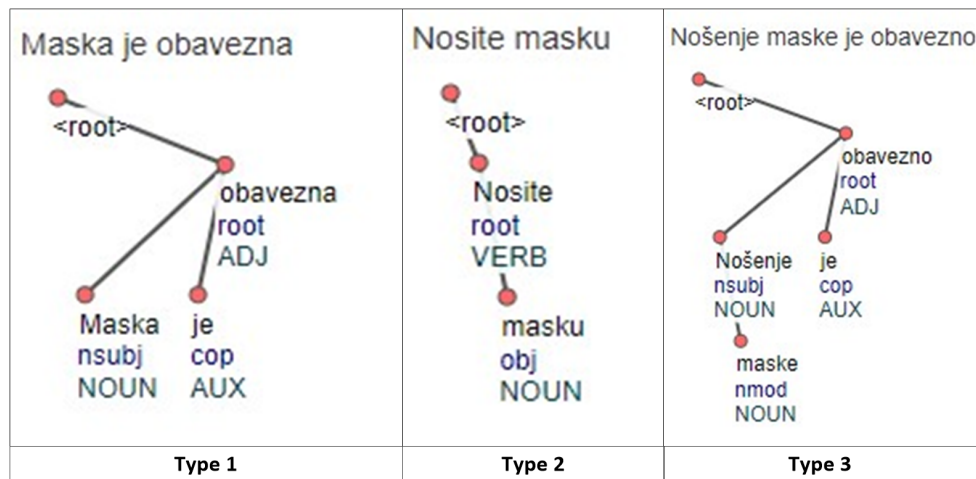


Figure 1: The basic syntactic structures of mask messages.

a specific way. *Osoba* and *lice* (person) are applicable to all types of premises, while visitors of *biblioteka* (library) and *sala* (hall) cannot be designated as *kupac*, *potrošač* or *mušterija*. Moreover, even though *mušterija*, *kupac* and *potrošač* can all be translated as ‘customer’, only *mušterija* is applicable to *menjačnica* (exchange office) and *frizerski salon* (hairdresser), while *kupac* and *potrošač* are not. On the other hand, the library patrons are customarily designated as *posetilac* (visitor) or *član* (member), while *mušterija* cannot be used. The larger and more versatile corpus is needed to fully investigate relations between designations for types of visitors and types of venues.

Finally, even in this small corpus we have ambiguity: *lice* can mean both face and person. In the former meaning it was used 9 times, e.g. *maska za lice* (face mask), while in the later case it was used 8 times.

#### 4 The Analysis of Mask Messages

In this section we will present basic syntactic patterns of one section of selected notices, namely the one related to wearing masks. There were 207 such statements, with a total of 1,466 tokens, and 1,365 words. The most frequent of 72 different nouns is *maska* (204), followed by *nošenje* (wearing) (91) and *ulazak* (entering) (37). The most frequent of 22 different verbs is *moliti* (to request) (29), followed by *nositi* (to wear) (25) and *staviti* (to put) (10). The most frequent of 21 different adjectives is *obavezan* (mandatory) (124), *zaštitni* (protective) (83) and *zatvoren* (enclosed) (8). In total, 8 different adverbs have been identified, of which only 2 occurred more than once: *obavezno* (mandatorily)

(14) and *isključivo* (exclusively) (2).

The most frequent noun *maska* is often characterized as *zaštitna maska* (protective mask), *higijenska maska* (hygienic mask) and *maska za lice* (face mask), and can be additionally described as *ličan* (personal). In the descriptions below [maska] stands for all these possibilities. Also, in this context *korišćenje* and *upotreba* (usage) are treated as synonyms for *nošenje* ([korišćenje]); similarly, synonym of *nositi* (to wear) is *koristiti* (to use), while *staviti* (to put) is also used in a similar context ([koristiti]).

As for their syntactic patterns, the majority of mask messages has one of 9 general forms discussed below. Their basic syntactic patterns were analyzed using the UDPipe (Straka and Straková, 2017) and they are presented in figures 1–3.<sup>3</sup> All these basic sentences can be modified with additional phrases: *Molimo Vas...* (Please...), *U ovom objektu...* (On this premises), *Pri ulasku u ...* (On entering in...), and *Svi kupci* (All customers...). For the recognition of these basic patterns with their various realizations we developed within Unitex/Gramlab<sup>4</sup> local grammars that are supported by Serbian morphological dictionaries (Stanković et al., 2021).

**1. [maska] je [obavezna]** – or “Mask is mandatory”. In this statement *maska* and *obavezan* can be replaced with synonyms (see Section 3), and the whole statement can be in plural: *maske su obavezne* (masks are mandatory). The auxiliary is sometimes omitted. This form was adopted by 25

<sup>3</sup>We analysed basic sentences using (Straka, 2020)

<sup>4</sup>Unitex/Gramlab – the Multilingual Corpus Processing Suite ([unitexgramlab.org](http://unitexgramlab.org))

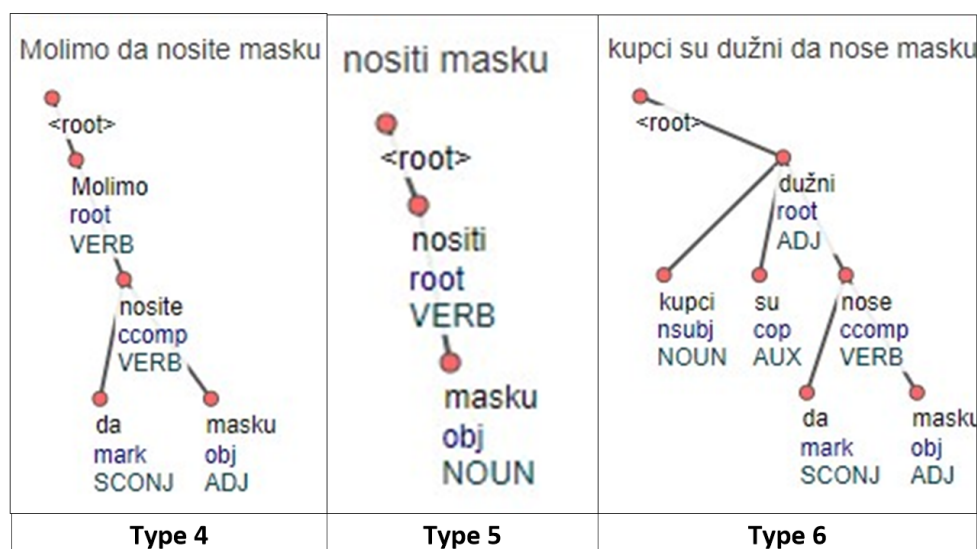


Figure 2: The basic syntactic structures of mask messages (continuation).

notices, two of which are: *Molimo Vas - maske su obavezne pri ulasku* (Please - masks are required upon entry) and *U lokalu obavezne zaštitne maske* (In premises protective masks mandatory).

**2. [nošenje] [maske] je [obavezno]** – or “Wearing a mask is mandatory”. Notices with this form occurred 99 times, and two of them are: *Nošenje maske u poslastičarnici je obavezno* (Wearing a mask in the pastry shop is mandatory) and *tokom boravka na fakultetu neophodno je nošenje maske* (it is necessary to wear a mask during your stay at the faculty premises). No message of this type was used with a phrase *Molimo Vas...* (Please...).

**3. [nosite] [masku]** – or “Wear a mask”. In this case verb *nositi* is in the imperative mood, 2nd person plural. There were 19 notices of this form, and two of them are: *Molimo, stavite masku pre ulaska u apoteku* (Please put on the mask before entering the pharmacy) and *Pri ulasku u radnju obavezno koristite masku* (Be sure to use a mask when entering the store).

**4. [Molimo da] [nosite] [masku]** – or “We entreat you to wear a mask”. In this case verb *nositi* is in the present tense, 2nd person plural. The structure remains the same if *Molimo da...* (We entreat you to...) is replaced by *Hvala što...* (Thank you for...). There were 15 notices of this type: *Hvala što nosite masku* (Thanks for wearing a mask) and *Molimo Vas da prilikom ulaska u prodajni prostor nosite zaštitnu masku* (Please when entering the shopping area wear a protective mask).

**5. [nositi] [masku]** – or “to wear a mask”. In this case verb *nositi* is in the infinitive. There were 2 notices of this type: *Masku staviti pre ulaska u agenciju* (Put on the mask before entering the agency) and *Obavezno koristiti zaštitnu masku* ((It is) Obligatory to use a protective mask).

**6. [kupci] su [dužni] da [nose] [masku]** – or “customers are required to wear a mask”. Here [dužni]={dužni, obavezni}, while [kupci] stands for all types of persons entering premises. In this case verb *nositi* is in the present tense, 3rd person plural. There were 5 notices of this type: *Kupci su obavezni da imaju zaštitnu masku* (Customers are required to have a protective mask) and *Sva lica dužna su da pri ulasku u objekat nose masku* (All persons are required to wear a mask when entering the facility).

**7. zabranjen je [ulaz] bez [maske]** – or “entry without a mask is prohibited”. Here [ulaz]={ulaz, ulazak, dolazak}. The auxiliary can be omitted. There were 9 notices of this type: *Strogo zabranjen ulaz bez maske* (entry without a mask is strictly prohibited) and *Ulaz u maloprodaju je zabranjen licima bez zaštitne maske!* (Entry into retail store is prohibited to persons without a protective mask). A variant of this structure is negated: **nije dozvoljen [ulaz] bez [maske]**, one of 4 retrieved examples is: *U knjižaru nije dozvoljen ulaz bez maske* (It is not allowed to enter the bookstore without a mask).

**8. [ulaz] je dozvoljen sa [maskom]** – or “entry allowed with a mask”. [ulaz] has the same values as before. The auxiliary can be omitted. There were

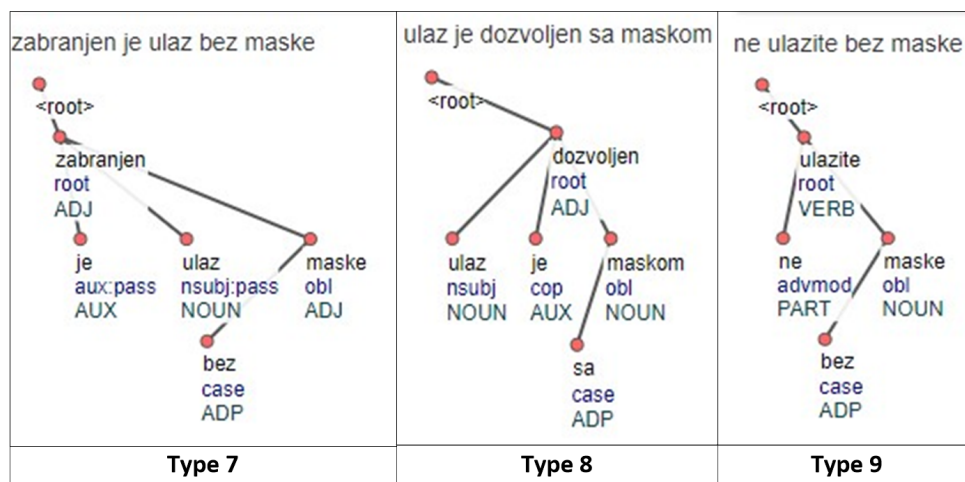


Figure 3: The basic syntactic structures of mask messages (continuation).

3 notices of this type: *ulazak je dozvoljen samo osobama sa zaštitnom maskom* (entry is allowed only to persons with a mask).

9. There were 7 notices that use the verb *ulaziti* (to enter). When used without negation these notices are similar to 4 – **Molimo da ulazite sa [maskom]** (Please enter with a mask). An example is *Molimo Vas da u galeriju ulazite s maskom* (Please enter the gallery with a mask). The form with the negation is **ne ulazite bez [maske]** (do not enter without a mask) where verb is in the imperative mood, 2nd person plural. An example is: *u radnju ne ulazite bez zaštitne maske* (Do not enter the shop without a protective mask). There were also examples with the verb in the infinitive (similar to 5) and in the 3rd person plural (similar to 6).

These 9 basic syntactic patterns describe 188 out of 207 messages about wearing a mask. Some of the remaining notices either do not use the word *maska* but a more general term, e.g. *Obavezno koristite mere zaštite od koronavirusa* (Be sure to use coronavirus protection measures), use some shortened expressions like *nošenje maske* (wearing a mask), or some specific form *Bez maske ne primamo u objekat* (We shall deny entry into the facility (to those) without a mask).

## 5 Hidden messages and sentiment values

Although the parts of notices related to mask wearing are very short, less than 7 words on average, some of them convey additional meaning to the main one, which is that one is requested to wear a mask. For instance, in two cases the statement contained *obavezno nošenje maski preko nosa i*

*usta* (mandatory to wear a mask over the nose and mouth) meaning that it is not enough to wear a mask, but also to wear it properly. Moreover, if notice said *ulaz u objekt s maskom* (entry in the facility (allowed only) with a mask), it would be possible to interpret it as if one was required to wear a mask only while entering the facility, but not throughout the visit. Therefore, some notices added explicit clarification: *Nije dozvoljen ulazak odnosno boravak lica bez ličnih zaštitnih maski* (It is not allowed for persons to enter or stay without personal protective masks).

The tone of messages related to masks vary across our corpus, and goes from very polite to severe or even unpleasant. Notes starting with *Molimo da...* or *Hvala Vam što...* always have a positive tone. However, some expressions like *Zabranjeno je...* (It is forbidden...) are never used with “please”, and cases like that were not found in our set. Also, a neutral expression like *Nošenje maske je obavezno* (Wearing a mask is obligatory) can be made more severe by adding *bez izuzetka* (without exceptions) or similar. The level of severity of mask notices according to their lexica and syntax can be ranked in the following way strating from those most strict:

- Messages using *zabranjeno je* (*Zabranjen ulazak bez maske* – Entry without a mask is forbidden);
- messages using imperative mood and/or negation (*Stavite masku* – Put a mask, *Ne ulazite bez maske* – Do not enter without mask);
- neutral messages (*Maska je obavezna* – Mask is obligatory);

- “Please” with the imperative mood and/or negation (*Molimo, stavite masku* – Please put a mask, *Molimo da ne ulazite bez maske* – Please, do not enter without a mask);
- “Please” addressing persons indirectly (*Molimo kupce da nose masku* – We ask customers to wear a mask);
- “Please” addressing persons directly (*Molimo Vas da nosite masku* – We ask you to wear a mask).

There is a Serbian proverb *Lepa reč gvozdена vrata otvara* (A nice word opens an iron gate). In this case, we cannot say which announcements, polite or strict, appealed to people more to respect the measures.

## 6 Generating mask messages

We used the local grammars developed for the recognition of 9 basic patterns with their various realizations that were systematized on the basis of data found in our corpus to generate possible mask messages within Unitex. In order to avoid an excessive number of possibilities, synonyms retrieved in our set, sometimes also hiponyms and hiperonyms, designating concepts *maska, kupac/osoba, objekat/radnja, ulazak* were not used for generation. By doing so we obtained:

**Type 1** – 19,520 messages, for example: *u ovoj prostoriji maska preko nosa i usta je obavezna* (in this room a mask over a nose and a mouth is obligatory);

**Type 2** – 10,944 messages, for example: *upotreba maski na licu u ovom objektu je obavezna* (the use of masks over the face in this facility is obligatory);

**Type 3** – 110,784 messages, for example: *molimo, stavite masku preko nosa i usta obavezno ako ulazite u ovu radnju* (please, put a mask over your nose and mouth obligatorily if you enter this shop);

**Type 4** – 62,160 messages, for example: *molimo da imate vašu masku na licu pri ulasku i za vreme boravka u ovoj prostoriji* (we ask you to have your mask on (your) face when entering and during (your) stay in this room);

**Type 5** – 14,240 messages, for example: *u objektu obavezno morate nositi masku* (in the facility you must be sure to wear a mask);

**Type 6** – 35,328 messages, for example: *mole se svi kupci da prilikom ulaska u ovu radnju obavezno stave masku* (all customers are asked to be sure to put a mask when entering this shop);

**Type 7** – 6,544 messages, for example: *ulazak nije dozvoljen osobama bez maske* (entering is not allowed to persons without a mask);

**Type 8** – 11,520 messages, for example: *dozvoljen je ulazak isključivo osobama sa maskom na licu* (only persons with a mask on (their) face are allowed to enter);

**Type 9** – 19,710 messages, for example *mole se kupci da ne ulaze u ovaj objekat bez maske* (the customers are asked not to enter this facility without a mask).

In a total, we produced 290,750 ready to use, correct messages all conveying the same basic meaning: “wear a mask”.

## 7 Conclusion

In this paper we presented the analysis of a set of notices collected from front doors of various premises that require compliance with protective measures against Covid-19. We analysed lexic and syntactic patterns of mask notices in more detail which enabled us to generate notices featuring one of their basic structures.

Our next step will be the production of paraphrased sentences with a full morphosyntactic description. Besides that we will analyse in the similar way messages about other Covid-19 protective measures. Moreover, we will collect other public announcements that emerge spontaneously and convey the similar meaning, like *Zatvarajte vrata za sobom* (close the door behind you) or *Ne primamo reklame* (We do not accept advertisements). The goal of our future project is to produce a big and versatile set of paraphrases.

## References

- Nisreen Naji Al-Khawaldeh, Imad Khawaldeh, Baker Bani-Khair, and Amal Al-Khawaldeh. 2017. An exploration of graffiti on university’s walls: A corpus-based discourse analysis study. *Indonesian Journal of Applied Linguistics*, 7(1):29–42.



- Alex Alonso. 1998. Urban graffiti on the city landscape. *San Diego State University*.
- Josie Bernicot, Olga Volckaert-Legrier, Antonine Goumi, and Alain Bert-Erboul. 2012. Forms and functions of SMS messages: A study of variations in a corpus written by adolescents. *Journal of Pragmatics*, 44(12):1701–1715.
- Jing Deng et al. 2010. Texting Christmas wishes in China: A view from pragmatics. *Bucharest Working Papers in Linguistics*, (2):115–127.
- Ilse Depraetere, Sofie Decock, and Nicolas Ruytenbeek. 2021. Linguistic (in)directness in Twitter complaints: A contrastive analysis of railway complaint interactions. *Journal of Pragmatics*, 171:215–233.
- Dejan Ivković. 2013. Pragmatics meets ideology: Digraphia and non-standard orthographic practices in Serbian online news forums. *Journal of Language and Politics*, 12(3):335–356.
- Gordana Jelić and Gordana Vekarić. 2019. Elliptical Constructions in SMS Communication. In *Značenje u jeziku – Od individualnog do kolektivnog. Zbornik radova sa međunarodnoga znanstvenog skupa Hrvatskoga društva za primenjenu lingvistiku održanoga od 16. do 18. svibnja 2019. u Rijeci*, pages 89–102.
- Danka Jokić, Ranka Stanković, Cvetana Krstev, and Branislava Šandrih. 2021. A Twitter Corpus and lexicon for abusive speech detection in Serbian. In *3rd Conference on Language, Data and Knowledge (LDK 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
- Adela Ljajić and Ulfeta Marovac. 2019. Improving sentiment analysis for Twitter data by handling negation rules in the Serbian language. *Computer Science and Information Systems*, 16(1):289–311.
- Marija S Milosavljević. 2018. Addressing the interlocutor in informative programmes. *Reči (Beograd)*, 10(1):15–29.
- Zorica Mršević. 2014. The beauty of graffiti. <http://zoricamrsevic.in.rs/english/GraffitiBeautyMrsevic4.pdf>.
- Biljana Ratković Njegovan, Maja Vukadinović, and Leposava Grubić Nešić. 2011. Characteristics and types of authority: the attitudes of young people. a case study. *Sociológia*, 43(6):657–673.
- Rachel Panckhurst. 2017. A digital corpus resource of authentic anonymized French text messages: 88milSMS—What about transcoding and linguistic annotation? *Digital Scholarship in the Humanities*, 32(suppl.1):i92–i102.
- Saša Petrović, Miles Osborne, and Victor Lavrenko. 2010. The Edinburgh Twitter corpus. In *Proceedings of the NAACL HLT 2010 workshop on computational linguistics in a world of social media*, pages 25–26.
- Vesna Polovina and Gordana Jelić. 2020. Shortening and Clipping in Serbian Text Messaging. *Anali Filološkog fakulteta*, 32(2):321–343. 18.
- Branislava Šandrih. 2018. Fingerprints in SMS Messages: Automatic Recognition of a Short Message Sender Using Gradient Boosting. In *3rd International Conference Computational Linguistics in Bulgaria (CLIB 2018)*. Department of Computational Linguistics at the Institute for Bulgarian Language with the Bulgarian Academy of Sciences: Sofia, Bulgaria, pages 203–210.
- Christoph Sauer. 2007. Christmas messages by heads of state. *Pragmatics & Beyond New Series (P&BNS)*, page 227.
- Katrin Schlund. 2014. Aspects of linguistic politeness in Serbian. A data-based comparison with German. *Linguistik online*, 69(7).
- Ranka Stanković, Cvetana Krstev, Rada Stijović, Mirjana Gočanin, and Mihailo Škorić. 2021. Towards Automatic Definition Extraction for Serbian. In *Proceedings of the XIX EURALEX Congress of the European Association for Lexicography: Lexicography for Inclusion (Volume 2)*, pages 695–704. Democritus University of Thrace.
- Milan Straka. 2020. UDPipe Croatian: Morphosyntactic Analysis of Raw Text. <https://live.european-language-grid.eu/catalogue/tool-service/437>.
- Milan Straka and Jana Straková. 2017. Tokenizing, POS tagging, lemmatizing and parsing UD 2.0 with UDPipe. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 88–99, Vancouver, Canada. Association for Computational Linguistics.
- Teresa Maria Włosowicz. 2011. Ways of expressing birthday, Christmas and New Year’s and Easter wishes in L2 and L3: Cross-cultural transfer and interlanguage pragmatics. In *Aspects of culture in second language acquisition and foreign language learning*, pages 217–231. Springer.