Glossos
Issue 6, Fall 2005
http://seelrc.org/glossos/
glossos@seelrc.org

Zrinka Jelaska & Milvia Gulešić Machata
University of Zagreb

PROTOTYPICALITY AND THE CONCEPT OF PHONEME

0. Introduction

The status of the cognitive model as a valid alternative to the classical models of linguistic categorization is supported by relatively few works on phonology (e.g. Taylor 2002, Nathan 1989, 1994). All phonological units (syllables, rhymes, nuclei, accents...), including phonemes, can be viewed as more or less prototypical, as well as can other linguistic representations. Taylor (2003) uses the analogy with polysemous categories to establish chain relationships on the basis of phonetic similarity between individual members of a phoneme, as well as to establish the central status of certain members of the category among different allophones, and the internal structure of the category of phoneme. He also shows that syllable structures can be represented by means of a network of phonological constructions with properties that are analogous to syntactic constructions (Taylor 2003, 263). This paper examines the principles that categorize sounds as phonemes in a language; in other words, it shows that phonemes of a given language do not have equal status: some phonemes are prototypical, whereas others are less prototypical, or (more or less) peripheral, with marginal phonemes beeing the most peripheral phonemes. Using the Croatian language as an example, the first part of the paper gives an overview of prototypical and more peripheral Croatian phonemes in order to determine what is involved in the concept of a phoneme. The second part of the paper presents examples from learners of Croatian as a second/foreign language (mainly Slovaks) that sheds additional light on factors involved in the conceptualization of phonemes.

1. The principles of phonemic categorization

Sounds are defined as phonemes on the basis of meaning. It is not the sound *per se*, but the

established meaning that gives the phoneme its status. But phonemes are nevertheless part of a word form. They are based on sounds and primarily realized as sounds.

1.1. Phonetic and phonological criteria

Phonemes as sounds are perceived and produced in speech. Therefore, their categorization depends on phonetic and phonological criteria, including pronunciation and hearing perception. Some clues to the prototypicality of sounds can be discovered by looking at the occurrence or hierarchical implications of phonemes in the languages of the world, the order of acquisition of sounds, and the similarity of sounds within a language. Sounds or pairs of sounds with clear pronunciation and hearing identity, such as dental *t* vs. velar *k*, or dental fricative *s* vs. palato-alveolar *š,* are preferred over sounds that are hard(er) to distinguish in hearing or production, such as lamino-postalveolar vs. lamino-prepalatal fricatives, prepalatal vs. palatal affricates, palatal vs. prepalatal stops, or pharyngeal *h* vs. glottal *h,* non-syllabic *i vs.* approximant *j*, and especially palatal plosive vs. palatal affricate, which do not seem to form a minimal pair in any language (e.g. Catford 1988).

Prototypical phonemes such as *a, i, u, p, k, s* appear in all or most of the languages of the world. Even within one language, they appear in many of prototypical words and distinguish between them, making minimal pairs in high-frequency words and other words that are viewed as prototypical words, such as Croatian *kit* 'whale', *kat* 'floor', *kut* 'corner', *kap* 'drop', *kup* 'pile', *kip* 'statue', etc. Hence, the prototypicality of phonemes is connected to the prototypicality of words they distinguish.

Due to the universal preference for open syllables, words such as *mala* 'little' are prototypical, words such as *vol* 'ox' are less prototypical. Words with conconant clusters such as *kruška* 'peach' are more peripheral than words without it, such as *kuša* '(he/she/it) tastes' or *ruka* 'hand'. Words such as *pokućstvo* 'furniture', because of the consonant cluster -*ćstv*-, are very peripheral in both Croatian and in general. Some words are more peripheral due to their choice of sounds, such as *gnjiljenje* 'decaying', which consists of the same and similar less prototypical sounds (palatal sonants *nj, lj, nj)* and *gniježđenje* 'nesting', which

is even more saturated by palatals.

1.2. Orthographic criteria

It is generally accepted that phonemic awareness of an illiterate child or adult is not the same as that of a literate person. Orthography can be confusing for phoneme categorization. Different words are judged as having different phonemes due to their spelling specificities. For example, when Croatian children first learn to write, they get introduced to letter *c,* representing sound *c* (dental affricate) in words such as *cipela* 'shoe', *maca* 'kitten'. Early on they write the name of the Croatian language and state with the letter 'c', based on production forms of those words: *hrvacki* 'Croatian', *Hrvacka* 'Croatia'. The same mistake is made by semi-illiterate speakers. They think that both words consist of seven sounds, e.g. phonemes /h r v a c k a/. However, the majority of literate Croatian speakers think the word *Hrvatska* consists of eight phonemes: /h r v a t s k a/, some think it has seven phonemes: /h r v a c k a/, a few are confused to start with and cannot decide if there is a phoneme *c* or the sequence of two phonemes *ts* between *a* and *k* (Jelaska 2004).

1.3. Morphological and (morpho)syntactic features

Words *ruka* 'hand', *noga* 'leg' and *svrha* 'purpose' are very frequent Croatian words, but in terms of their morphology or morphophonology, they are more peripheral than less frequent words such as *kuka* 'hook', *kuga* 'plague' and *ploha* 'surface'. The former undergo sibilarization in the dative and locative case (*ruci, nozi, svrsi*), and the latter do not. In other words, phonemes *k, g* and *h* that undergo sibilarization are somewhat less prototypical than the ones that do not change in declension and conjugation.

Morphosyntactic criteria are also important: phonemes from words with the same morphological form are closer to the prototype than those from words with different morphological forms. There is a constant debate among phonologists who recognize the phonemic status of the diphthong *ie* when the minimal pair proving its existence comes from the genitive singular *dijela* [die:la:] from *dio* 'part' vs. the genitive plural such as *djela*

[dje:la:] and those who don't agree with this theory. The same can be applied to syntactic forms. The phonemes with a distinctive function within the same word class are viewed as more prototypical than those having minimal pairs from different word classes (cf. Brozović 1991).

1.4. Lexical criteria

One set of criteria for prototypicality depends on the words that give phonemes their status, e.g. the status of words in which they appear. Prototypical phonemes appear in prototypical words. Prototypical words are common, simple words with prototypical forms and meanings (Jelaska 2004), and frequent words used in everyday conversation or in discussions on general topics, which are at the same time old Croatian words of Slavic origin. Therefore, some phonemes are less prototypical or peripheral because they distinguish mostly or only peripheral words. A word is less prototypical if it is marked by any of the mentioned or other criteria. The more a word is marked, less prototypical it is. Some words are less prototypical because they are less frequent (e.g. *spavačica* 'female sleeper', *hvat* 'fathom'). Others are less prototypical because they belong to marked idioms: dialectal, colloquial (e.g. *dućan* 'store', *uhapsiti* 'bust '), slang (e.g. *škvadra* 'guys'), professional jargon (e.g. *ksenofobija* ′xenophobia′, *aneks* 'annex'), or technical terminology (e.g. *šumnik* 'true consonant, obstruent'), etc. Others are very peripheral because of their form, written and/or spoken: abbreviations such as *HT*, loan words such as *restaurant*, etc.

One marginal Croatian phoneme is the syllabic dental nasal in scientific internationalisms such as *njutn*, and barbarisms or exonymes in cooking jargon such as *šmarn* - at least this is the opinion of those who believe that syllabic consonants are different phonemes than their non-syllabic pairs. The voiced dental affricate *dz* in the music term *mezzosopran* has a *s*imilar marginal phonemic status. Some speakers recognize it as sound/phoneme *dz*, others as phoneme *c* because *dz* is its allophone in Croatian, and some as *z*, the only voiced dental sibilant (this is supported by the letters zz). The same is or was the case with foreign name *Ševernadze*. Sounds that appear in a few foreign names of people, nations, cities,

newspapers, etc. are peripheral, to the extent that many speakers do not even consider them phonemes, hence marginal, such are the central vowel ə in the article in *Le Corbusier* and *Le Matin* (Škarić 1991), and the syllabic palatal lateral in *Kremlj*.

The prototypical status of phonological units with regard to the status of words is observable at other word levels, too. For example, the consonant cluster *tk* (dental+velar voiceless occlusive) is more typical in the syllable onset than *kt* (velar+dental occlusive). The former appears in few Croatian words that are otherwise very frequent and simple, such as *tko* 'who' and '*tkanina*' 'cloth', while *kt* appears in the linguistic term *ktetik* 'possesive adjective formed after a place name'. The status of their voiced pairs is exactly the opposite: *gd* (velar+dental voiced occlusive) is much closer to the prototypical cluster because it appears in a frequent and simple Croatian everyday word *gdje* 'where' and its derivations, whereas *dg* appears only in the dictionary form *dgunja* as adaptation from Greek *kydonion melo*s (the standard form is *dunja* 'quince'). To take another phonological example that does not concern phonemes, the co-occurrence of the high tone and stress in the same non-initial syllable is marginal, although not that rare, because it appears in lexically peripheral words such as abbreviation *HBK* - *habek**a***, loan words such as *buffet* - *bif**e***, international names such as *Ant**o**nija*, *Mir**e**la*, compounds such as *vele**i**zdaja* 'high treason', or morphologically peripheral words such as the genitive plural *tren**u**taka* 'of moments', or *računala* 'of computers'.

1.5. Psycholinguistic and sociolinguistic criteria

As it was stated earlier, prototypical words are considered to be Croatian words of Slavic origin. These are so called *domaće riječi* 'domestic words'. That means that borrowings are non-prototypical. Recognizable borrowings are more peripheral than only historically or sociolinguistically marked borrowings. For example, if sound sequences violate syllabic template of Croatian language, their foreign origin is evident from the form. Croatian prototypical words have only one consonant in a coda, only *-st, -št, -zd* and *-žd* appear as consonant clusters at the end of words. Therefore, borrowings such as *farinks* 'pharynx' and *ktetik* are recognized as words of foreign origin, or at least as peripheral words due to

their form. The same applies to Slavic borrowings with non-typical consonant clusters such as -ršč in *boršč*, that are also noticeably foreign. Slavic words that do not violate phonological constrains are nevertheless sometimes recognizable by their morphological form, e.g. *golubaja*. Croatian speakers do not have to know the language it comes from, but they can still guess it as a word of some other Slavic language rather than Croatian. They recognize and understand the root '*golub-'* because it is the same as the Croatian word *golub* 'pigeon', and therefore presume the rest is another Slavic morpheme or a few of them.

Some linguists still consider the status of *dž* in Croatian questionable, since it appears almost exclusively in words of Turkish (e.g. *hodža, džezva*) and English (e.g. *džungla, džemper*) origin (e.g. Silić 1992). Dilemmas and controversies about the phonemic form of Hungary and Hungarians: '*Mađar'* or '*Madžar'* are to some extent caused by this, as the first is borrowed directly from Hungarian, while the second is transmitted through Turkish.

Apart from the above-mentioned approach to borrowings that do not violate Croatian phonologic or morphologic constrains, there are other sociolinguistic and psycholinguistic factors that could also make a phoneme less prototypical. One psycholinguistic factor is the ease of acquisition among the native speakers, which includes speakers of different dialects. For example, standard Croatian has two sets of affricate palatals: palato-alveolar (voiceless *č* and voiced *dž*) and palatal or prepalatal (voiceless *ć* and voiced *đ*). There are many speakers of Croatian dialects with only one (voiced and voiceless) pair of palatal affricates: post-alveolar, or palato-alveolar, or palatal. Since the Croatian standard is to some extent L2 for those speakers (they would be called bidialectal, or simply bilingual), in communication they resort to means other than palato-alveolar vs. (pre)palatal distinction. The sociolinguistic status of speakers who distinguish both sets of affricates is not as high as it used to be, and therefore the distinction is preserved in writing, though there are occasional disputes about their phonemic status in scholarly journals or magazines. In speech, some speakers of standard language insist on keeping the two-phonemic system, some do not care, and others would like to make it one-phonemic.

1.6. Combinations of criteria

However, the listed criteria may play a role simultaneously. Some words and their phonological units could be judged as more prototypical by some criteria and less prototypical by other. For example, Croatian words such as *ptica* 'bird, *pčela* 'bee', *tko* 'who', *gdje* 'where' are prototypical in terms of meaning and frequency, but their forms are less prototypical (or more peripheral) than those of less frequent words such as *ticalo* 'feeler', *čedo* 'infant', *kolo* 'wheel dance' *djetao* 'woodpecker', *pir* 'feast', *pelud* 'pollen', *tokar* 'turner', *gegavac* 'waddler' etc. Their outcome offers a variety of relationships between phonemes just as any other concept. For examples, in debates on palatal affricates distinction, only *č* and *ć* are mentioned, and *dž* and *đ* are almost never mentioned, which shows that they do not have the same status in terms of prototypicality.

According to the number of criteria for phoneme prototypicality in Croatian, and their relationship, it could be expected that the phonemes that are marginal due to several different criteria will not only cause trouble to native speakers and cause disagreements between scholars, but will sometimes not even be listed among phonemes (Jelaska 2004). This is precisely the case. For example, the phonemic and graphemic form of the word *vrabac* G *vrapca* 'sparrow ' do not cause problems, because *vrabac* is a frequent word, but *kobac* G *kopca* 'sparrow-hawk' does because *kobac* is a low frequency word. Words like *mladac* G *mladca* 'whippersnapper', which is stylistically marked, and *mlatac* G *mlatca* 'flail; swingle', which is very rare, are a real bone of contention among both linguists and other speakers. Hence, the categories of phonemes in a language have fuzzy edges, just like other categories.

2. Evidence from the CSL

As it is not always convenient to conduct research on illiterate adults, a more available group of speakers with distinctive phoneme conceptualization are learners of Croatian as a foreign language.

2.1. Phonemes and sounds

The SLA research points to three different groups of sounds that differ between languages: the same/equal sounds, similar sounds and different sounds. Contrastive approach predicts that foreign language learners will have difficulty acquiring new sounds/phonemes as well as phonemes realized by different sounds in their mother tongue. (As can be seen, it is sometimes hard to distinguish between phonemes and their realizations by sounds when we speak about this issue in general.) Foreigners will tend to transfer sounds from their mother tongue. To some extent, this prediction is true. Learners often have no problems with sounds that are the same as in their mother tongue, but they do have problems with sounds that are different or non-existing in their mother tongue. For example, Hungarians, Americans and many other learners of Croatian find it difficult to pronounce Croatian *l'* - they pronounce the sequence *l*+*j* instead of a lateral palatal. Chinese speakers have difficulties in pronouncing Croatian trill *r*, Hungarian and German speakers have difficulties in pronouncing Croatian voiced palato-alveolar affricate *dž*, and Spanish speakers have problems with pronouncing Croatian voiceless dental fricative *s* and so on.

However, sometimes different phonemes do not cause problems. Developmental factors based on prototypicality account for the fact that, in foreign language acquisition of phonemes, foreigners do not have problems with some new phonemes, e.g. with Croatian *j* if *j* does not have phonemic status in a foreign language.

It is not only the different sounds that can cause difficulties with pronunciation. Sometimes problems arise with similar sounds - sounds which are not the same, but in a way similar in two or more languages. In this case, learners are often faced with the transfer problem. For example, Croatian *r* as an apico-alveolar trill presents a problem for native speakers of English, German and French. They replace it with their respectful r-sounds: taps, uvular trill, uvular fricative or even a uvular approximant. Since this practice does not cause a communicational problem and the phonemic role of *r* is recognized, it is only their allophonic, i.e. orthoepic form that is not acquired. Hence, they have only identification

problems: they are readily recognized by Croatian speakers not only as foreigners, but often as American, German and French speakers.

In contrast, Chinese speakers have a communicational problem with Croatian *r* as they often do not acquire it as a phoneme and have difficulties producing any sound that would be recognized as *r* by native Croatian speakers. As with Japanese speakers, they sometimes replace *r* with lateral *l*, which encroaches on the space of another Croatian phoneme. This shows that the category of a phoneme in a language can include many peripheral, even some marginal members produced by foreigners unless they invade the space of another phoneme.

An interesting example of sound perception and phonemic categorization comes from Slovak speakers acquiring Croatian velar fricative. The Slovak language has two velar fricative phonemes: one is voiceless, as in Croatian, and the other is voiced. The contrast between voiced and voiceless sounds is very important in Croatian, although there are three voiceless phonemes that have no voiced counterparts: *f, c, h*. Their voiced pairs play a role of allophones. While Croatians would notice when e.g. Italian speakers replace voiceless *c* by *dz* in words such as *dzuri* instead of *curi*, and sometimes even misunderstand it, they do not notice that something is wrong when a Slovak who speaks Croatian uses Slovak voiced instead of Croatian voiceless velar fricative! This can, in part, be explained by the phenomenon of language economy – a similar sound is considered to be appropriate since it is similar and, more importantly, cannot cause a communication problem. We already mentioned that other "similar sounds" were noticed and attributed to an identification problem.

2.2. The role of the written form

The role of written language in establishing phonemes and the mental image of a word responsible for pronunciation become more evident in the case of Croatian SL learners who cannot pronounce certain Croatian words because of their written form. The most interesting example comes from a girl who was drinking coffee without milk for days

because she could not order that in Croatian language (the case was presented by L. Cvikić, personal communication). The girl learned that she should use words *kavu s mlijekom*, which she was introduced simultaneously in the written and spoken form. It was the written picture that was responsible for the trouble, since the written form is a part of phoneme categorization, at least at the conscious level. The trigraph *ije* for diphtong *ie* was causing the pronunciation problem. Only when the teacher finally wrote *mliekom* to help her was she able to order and drink coffee with milk.

The experience of different teachers of CSL does not involve the case of an equal sound causing acquisition problems, with one major exception concerning Slovak speakers. It was discussed in Gulešić (2003), but we will explain it here at length because it is most unusual, or even unexpected, language mistake and because the original article is in Croatian.

Just like Americans, Germans, and Hungarians, who have difficulties in producing a Croatian lateral palatal because they lack such a phoneme in their language, Slovak learners often have the same problem with this sound. This is surprising because Slovak has the same phoneme realized by the same sound. The reason for this is the sound's graphemic form. In Croatian, the lateral palatal is written by a digraph *lj*, while in Slovak it is written by a monograph *ľ*. This is why Slovak learners of Croatian often replace the lateral palatal by two sounds, eg.*[lį̇udi], *[lį̇ubav]. The proof that we are dealing with visual, and not developmental influence, comes from the needed means of learning. While the Hungarians, English, and Germans need phonetic classes to acquire Croatian phonemes, Slovaks need only to master the different orthography. They can immediately pronounce Croatian words with /Ĺ/ if they hear it as a sound or if those words are written in Slovak orthography, e.g. *ľudi, ľubav* rather than Croatian *ljudi, ljubav*.

A similar problem of orthography interference can be found in Croatian words in which vowels *i* or *e* follow dental or alveolar sounds *d, t, n, l*. Slovak orthography does not mark palatalized pronunciation of consonants *d, t, n, l* in front of *e, i, í* - they write *neskoro*, and pronounce [ńeskoro]. Slovak speakers transfer this rule and pronounce Croatian *učionica*

as *[učiońica] instead of [učionica]. This is prevented by writing *učionica* with the Slovak grapheme *y*: *učionyca*, which represents the same vowel as *i* in Croatian and Slovak, but prevents palatalization of the preceding consonant. Slovak students will produce the lexeme *učionyca* with the Croatian dental nasal phoneme, and *učionica* with the palatal nasal.

Sometimes letters from a third language are used to solve the problem of phonemic or allophonic form. For example, Slovak speakers often pronounce Croatian words *krv* 'blood', and *ovca* 'sheep' as [kr̩u̯] and [ou̯ca], transferring Slovak phonetic rule about syllable-final *v* that is replaced by [u̯]. As Croatian and Slovak words have the same spelling, some other graphem is needed, such as W to represent [v], *KrW* and *oWca* will be written instead of *krv* and *ovca* to prevent the application of the Slovak rule. A similar transfer is found among vowels. As Slovaks mark long vowels in writing and do not mark short ones, they produce all Croatian vowels in reading as short, which influences their speech production.

2.3. Different treatment of phonemic errors

Croatian speakers treat phonemic replacements in their language by foreigners differently. Some are viewed as phonemic replacement (e.g. *nj* instead of *n*), some as pronunciation (allophonic) replacement (e.g. ou̯ca instead of ovca because u̯ has no phonemic status), and some are not even noticed (e.g. voiced *h* instead of voiceles *x)*.

It is quite understandable that native speakers notice phonemic replacements, although those replacements do no necessarily cause communication problems. For example, when foreigners say *učionjica* instead of *učionica*, Croatian speakers will recognize the word *učionica* because there is no similar Croatian word. But if foreigners pronounce the name *Anji* instead of *Ani* when two women *Ana* and *Anja* could be possible subjects, communication would fail. Is is not quite clear that native speakers should notice alophonic replacements at all, at least consciously, but if they do, they should treat them in a similar manner. But they do not. They consider replacement of voiceless *h* by a voiced sound no

problem at all, they even have difficulty noticing it. They cosider replacement of alveolar trill by tap or uvular trill just an identification problem. They consider replacement of voiceless dental affricate *c* by voiced *dz* as a communicational and identification problem. Why do Croatian speakers consider those replacements differently? Remember that none of these three (or four) sounds has a phonemic role.

The prototypicality principle gives a possible explanation. First, dental place of articulation is more prototypical as it offers possibility for more contrasting sounds than the velar one. Second, voiced dental affricate has a status of marginal phoneme due to the dialectal phonemic status (e.g. in Međimurje, Dubrovnik). In some borrowings (Italian music term, Ukrainian family name), even a digraph *dz* would be a pair of grapheme *c*, as *dž* is a graphemic pair of grapheme *č*, while the voiced velar fricative lacks ready-made grapheme.

3. Conclusion

Phonemes are categories organized around prototypical members not only within themselves, but between themselves. The less prototypical and more marginal phonemes differ from the central phonemes on the basis of different criteria, some of which were listed in this paper to show what is involved in the concept of phoneme. This improves our understanding of why was it so complicated for some generative linguists to capture phonemes within their theoretical framework, which made them reluctant to even include phonemes among (regular) linguistic units. The prototypicality principle offers an appropriate theoretical framework of phonemes and their status within the language, gives an explanation for many disputes among theoretical linguists on the subject of phonemes, helps explain foreign phoneme acquisition and can also give some suggestions for acquiring and teaching foreign phonemes.

4. References

Brozović, D. (1991) "Fonologija hrvatskoga književnog jezika." In: S. Babić et al., *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika*, Zagreb, Hrvatska akademija znanosti i umjetnosti i Globus, 379-452.

Catford, J. C. (1988) *A practical Introduction to Phonetics*, Clarendon Press, Oxford.

Gulešić, M. (2003) "Srodnost dvaju jezičnih sustava – prednost ili/i nedostatak u usvajanju jezika." In Stolac, D. et al. (eds.), *Psiholingvistika i kognitivna znanost u hrvatskoj primijenjenoj lingvistici,* Zagreb i Rijeka, 289-301.

Jelaska, Z. (2004) *Fonološki opisi hrvatskoga jezika*. Zagreb, Hrvatska sveučilišna naklada.

Nathan G.S. (1989) "Preliminaries to a theory of phonological substance: The substance of sonority." In Corrigan et al. (eds.), *Linguistic Categorization.* Amsterdam, Benjamin, 55-67.

Nathan G.S. (1994) "How the phoneme inventory gets its shape: Cognitive grammar's views of phonological systems." *Rivista di Linguistica* 6, 275-87.

Silić, J. (1992) "Status skupova *st* i *žd* u hrvatskom jeziku." *Suvremena lingvistika* 8/2 (34/9, 263-280.

Škarić, I. (1991) "Fonetika hrvatskoga književnog jezika." In: S. Babić et al., *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika*, Zagreb, Hrvatska akademija znanosti i umjetnosti i Globus, 61-378.

Taylor, J. (2002) *Cognitive Grammar*. Oxford, Oxford University Press.

Taylor, J. (2003) *Linguistic Categorization.* Oxford, Oxford University Press.