

**Vlasta Erdeljac, Anabela Lendić, Martina Sekulić Sović**  
*Sveučilište u Zagrebu, Filozofski fakultet*

### **Prikaz baze psiholingvističkih parametara riječi u hrvatskom jeziku**

Dosadašnja psiholingvistička i neurolingvistička istraživanja pokazala su da različite leksičke varijable u znatnoj mjeri utječu na točnost i brzinu kojima se riječi u upotrebi (prepoznavanju ili proizvodnji) prizivaju iz pamćenja, imenuju i klasificiraju. Empirijski prikupljeni podaci predstavljaju s jedne strane temelj za psiholingvističke i neurolingvističke teorijske pretpostavke, a s druge strane su i osnovni kriterij valorizacije tih pretpostavki. U pripremi materijala za bihevioralna jezična istraživanja te analizi dobivenih rezultata nužno je kontrolirati veliki broj različitih kognitivno i jezično utemeljenih varijabli, stoga je dostupnost normi za subjektivne leksičke varijable u hrvatskom jeziku jedan od neupitnih istraživačkih preduvjeta za sva istraživanja tog tipa. Osnovni je cilj ovog rada prikazati novi resurs za psiho- i neurolingvistička istraživanja – leksičku bazu riječi hrvatskog jezika opisanih s nekoliko temeljnih psiholingvističkih parametara: subjektivnu frekventnost riječi u upotrebi, predočivost, apstraktnost/konkretnost, poznatost, dob usvajanja riječi te asocijacijsku povezanost riječi. Psiholingvističko potpodručje asocijativnih odnosa u mentalnom leksikonu istraživački interes koncentrira na odnos između leksičkih i konceptualnih struktura. Istraživačka pitanja koja problematiziraju psiholingvističke varijable i njihovu ulogu u organizaciji i načelima funkcioniranja mentalnog leksikona mogu se promatrati u kontekstu povezivanja psiholingvistike, teorijski i eksperimentalno, s drugim lingvističkim i nelingvističkim disciplinama (prije svega semantikom i informacijskim znanostima). Iako još u ranim fazama primjene, razni informacijski alati uspješno su implementirani u psiholingvističkim istraživanjima za engleski jezik (AMT, Ibox farm, psiTurk i LingoTurk). U hrvatskom jeziku takva su istraživanja izuzetno rijetka. Ovakva vrsta interdisciplinarnog istraživačkog pristupa temelji se na primjeni informacijskih sustava i alata potpomognutih radom mnoštva koji kao osnovnu pretpostavku ističu sudjelovanje velikog broja korisnika i raznolikost njihovih mišljenja u jezičnim istraživanjima. Pri tome se postavlja pitanje mogu li se i u kojoj mjeri informacijski alati osmišljeni za engleski jezik prilagoditi i unaprijediti za primjenu u psiholingvističkim istraživanjima hrvatskog jezika. U preliminarnom istraživanju za hrvatski jezik bez primjene informacijskih alata testirano je šest parametara na stotinu zdravih ispitanika u dobi od 20 do 30 godina. Ispitanici su procjenjivali riječi putem skalarne procjene na ljestvici od jedan do pet. U daljnjem razvijanju baze, na temelju empirijskih istraživanja i analiza povećavat će se broj ispitanih varijabli i broj ispitanika uz pomoć novih informacijskih alata. Izrada baze psiholingvističkih parametara doprinos je u pružanju pouzdanih podataka na kojima se grade teorije i modeli jezičnog procesiranja, ali i instrumentarij za jezičnu dijagnostiku i rehabilitaciju u kliničkoj praksi.

### **Introducing the psycholinguistic variable database of the words in the Croatian language**

Psycholinguistic and neurolinguistic research has shown that various lexical variables have a significant influence on the accuracy and speed-rate of (in both language reception/perception and production) retrieving words from the mental lexicon, using words in naming, and classifying words. Empirically collected data represent, on the one hand, the basis for psycholinguistic and neurolinguistic theoretical assumptions, and on the other hand, are the basic criterion for the valorization of these assumptions. In preparing materials for behavioral linguistic research and the analysis of the obtained results, it is necessary to control a large number of different cognitive and linguistic variables, therefore highlighting the availability of norms for subjective lexical variables in Croatian as one of the crucial research preconditions for conducting all research of this type. The main objective of this paper is to present a new resource for psycholinguistics and neurolinguistic research – the lexical database consisting of the words of the Croatian language described by using several key psycholinguistic parameters: subjective frequency of the words, imageability, abstractness/concreteness, familiarity, age of acquisition of the words, and associative relations of the words. The psycholinguistic subfield of associative relations in the mental lexicon explores the relationship between lexical and conceptual structures. Research questions problematizing psycholinguistic variables and their role in the organization and principles of the functioning of the mental lexicon can be observed in the context of relating psycholinguistics, both theoretically and experimentally, with other linguistic and non-linguistic disciplines (primarily semantics and information sciences). Although still in the early stages of the application, various information tools have been successfully implemented in the psycholinguistic research for the English language (AMT, Ibox farm, psiTurk and LingoTurk). In the Croatian language, such research has been rarely conducted. This type of interdisciplinary research approach is based on implementing information systems and tools supported by crowdsourcing, which as its basic assumption emphasizes the participation of a large number of users and the diversity of their opinion in linguistic research. However, this raises the question of whether and to what extent information tools designed for English can be adapted and improved for use in the psycholinguistic research for the Croatian language. In the preliminary study for the Croatian language, without the use of information tools, six parameters were tested on one hundred healthy subjects at the age of 20 to 30 years. Participants evaluated the words using scalar estimations on a scale of one to five. In the further development of the database, based on empirical research and analysis, the number of variables tested and the number of participants will be increased using new information tools. Creating a database of psycholinguistic parameters contributes to providing reliable data on which theories and models of language processing are developed, as well as tools for language diagnostics and rehabilitation in clinical practice.