

SRoL - Web-based Resources and Tools used for e-Learning of Languages and Language Technology

Silvia Monica Feraru¹, Horia-Nicolai Teodorescu^{1,2}

(1) Institute for Computer Science, Romanian Academy, Bd. Carol I nr. 8, Iași, România

(2) Gheorghe Asachi Technical University of Iași, Iași, România

E-mail: mferaru@etc.tuiasi.ro

Abstract

The SRoL Web-based spoken language repository and tool collection was developed during several years by the collaboration of groups from the Institute for Computer Science of the Romanian Academy, CERFS Excellence Center in "Gheorghe Asachi" Technical University of Iasi and staff of the discipline of Language Technology, Computer Science Faculty, "Al.I. Cuza" University. The web site includes thousands of voice recordings grouped on sections like "Basic sounds of the Romanian language", "Emotional voices", "Specific language processes", "Pathological voices", "Comparison of natural and synthetic speech", "Gnathophonics and Gnathosonics". The recordings are annotated and documented according to proprietary methodology and protocols. Moreover, we included on the site extended documentation on the Romanian language, speech technology, and tools produced by us, for voice analysis. The resources are a part of the CLARIN European Network for Language Resources. The resources and tools are useful in virtual learning for phonetics of the Romanian language, speech technology and medical subjects related to voice. We report on several applications in language learning and voice technology classes.

Keywords: spoken language resources, voice education, speech and language, "dictionary of sounds", educational and research purposes

1 Introduction

In a world where the Web / Internet communication is pervasive, computer is more than a study topic for everyone, it is a ubiquitous tool. Computers serve for more than doing computations, they are now one of most used means of communication and interaction – the very basis of any educational system. As a consequence, computer-based education is an obvious choice whenever a distance separates the learner and the learning person. In a general sense, computer-based education and virtual education based on Internet is today an undeniable fact of life in every academic campus. While computers and the network are the means, the spoken language represents the prevalent support of communication in the teaching-learning process. Hence the natural need to address e-learning and virtual learning of languages, voice and phonetics, voice pathology, and other aspects related to voice and spoken language.

In view of the above, we built during a timeframe of about five years a web site that offers the possibility of teaching and learning various aspects on the Romanian language

based on an annotated corpus freely accessible on the Internet. The corpus is complemented with in-depth phonetic and linguistic analyses, moreover with specific tools accessible by users from everywhere through the web (Zbancioc, M., 2006), (Teodorescu et al., 2006), (Teodorescu et al., 2007b), (Teodorescu H.N., Feraru M., 2007), (Teodorescu H.N., Feraru M., 2008). This instrument has a high level of dimensionality and aims to numerous aspects of the language that are not typical features in language corpora. This makes this "corpus-tool" the unique instrument of its kind existing today in the domain (Teodorescu et al., 2007a).

Since the second author initiated, five years ago, the Project "The Sounds of Romanian Language" (SRoL), the team increased to 8 researches. During the recent years, we studied different emotional speech database which can help in education and re-education of speech, in diagnosis and treatment, in learning a language aided by computer; examples of results published are (Teodorescu H.N. et al., 2007a), (Feraru M., Teodorescu H.N., 2008) etc.

Voice and language e-education is a topic addressed by many research and educational groups. Solomon studied the possibilities and issues of learning with and about computers in schools or in any other learning environment (Solomon C., 1988). The *Eric Education Resources Page* shows the importance of computer assisted education of speech and voice (Wise, B.W., Olson, R.K., 1994). On the other side, web-based educational resources and training have received attention during the last decade. Åke Olofsson offers a simple method of compensation for word decoding problems, by having the micro-computer which pronounces the words which can not be read. Olofsson uses a program developed for the IBM-PC/AT and a Scandinavian multilingual text-to-speech unit, and children can read a textfile on the monitor, and use a mouse to request the immediate pronunciation of a word (Olofsson Å., 1992).

The computer-assisted learning language software helps the interaction between student and computer by speech, by sound effects, by animation, by video, not only text. On the other hand, these are restricted by the mouse and keyboard, hence it is necessary an active interaction by spoken language through computer (Cameron K., 1999). Speech recognition offers the possibilities to computer-assisted learning language to have an active participation by oral reading and conversation. CALL system has recordings spelling by the native speaker. The user compares the quality of her pronunciation with model recordings.

In another direction of research, Warschaue observes the uses of online communications for language teaching. He determined that the interest in this domain grows day by day. He proposed a conceptual framework for understanding the role of the interaction assisted by computer (Warschaue M., 1997). Lundberg considers the computer a tool of remediation in the education of students with reading disabilities as dyslexic students which can benefit by computer training in correct reading and spelling the words (Lundberg I., 1995).

A speech database is a collection of files with sounds, structured after its own purpose. The SRoL resource (corpus) is located at the address (http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/index.htm). The initiator conceived SRoL as an Internet-based "dictionary of sounds and words" for the Romanian language supplemented with specific manifestations of voice (including pathologies) and various tools. The SRoL

database includes files with vowels, consonants, diphthongs, sentences with emotional states, linguistic particularities for the Romanian language, dialectal voices, and gnathosonic and gnathophonic sounds. It is the first Internet based annotated database of emotional speech for the Romanian language which contains more than 1500 recordings in different coding formats (wav, ogg, txt / 22 kHz/ 24bit/ 16 bits). The phonetic recordings which refer to an annotated emotional speech corpus (database) are registered to ORDA. Figure 1 illustrates the home page of the SRoL speech database, which has English and French versions as well.



Figure 1. The frontpage of the SRoL project on the web, at http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/index.htm.

In this paper, we provide details about the applications of this database and about the SRoL-web database, available to the address http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/index.htm.

2 Applications for learning Romanian language

One of the goals of SRoL the web site is to provide a free Romanian database for students and researchers, for linguists, for teachers, in view of teaching, learning and analysis the Romanian language sounds. The database includes the pronunciation corpus and related documentation. The database contains among others, sections with:

- recordings of syllables and words pronounced in various contexts, like accentuated word, interrogative sentences, exclamations, various emotions conveyed by the speaker, etc. This part of the database is aimed as a source for concatenative synthesizers and as benchmark for the voice recognition systems – isolated words;
- files of sounds, syllables and words pronounced by persons with various pathologies; this section may be useful in medical and phonological researches;
- files with professional voices (“perfect” pronunciations), as well as non-professional voices, the “voices of the people in the street”. For the moment, we concentrate on voices from the Iaşi region (East Romania) and middle area of Moldova.

Learning and teaching languages require well documented audio-visual tools that exemplify and fully explain spelling for a large variety of voices and contextual and emotional states. While former methods, like tape recordings and audio disks have been

helpful, the multimedia Internet-based tools offer tremendously increased capabilities. SRoL represents such a tool for the Romanian language. Not only it is the first for the Romanian language, but its multidimensionality makes it somewhat unique and novel in concept for language learning and teaching in general.

Consider the case of a foreign student who wants to improve her Romanian pronunciation by comparing the prosody of her voice with the prosody of native speakers. The student utters a sentence (from those included in the site), then opens Wasp or another similar tool and displays the energy and fundamental frequency in her voice. She then compares these prosodic features to the ones of native speakers and tries to improve her prosody until she produces correct prosodic patterns. Also, the student can compare formant values and try improving the formants of the vowels she pronounces.

This instrument is useful for learning to improve communication, moreover for human-computer speech interaction, for security, for medical applications, for video-games and interactive TV, for teachers, in the study of the Romanian language, etc.

3 Applications in medical education and re-education of speech

Voice education is needed whenever a voice pathology including some neurologic and psychiatric disorders or pathology of the vocal tract occurs. Several groups have addressed the voice re-education topic (Lundberg I., 1995), (Olofsson Å., 1992).

Till now, we included in SRoL words pronounced by persons with minor pathologies as trembling voice. We have demonstrated in our research that splitting the signal in frequency bands that correspond to the peak of F0 – F1 formants and respectively to the peak of F2 – F3 formants helps improving the discrimination process in a significant way. The use of fractal dimensions in assessing the jitter or shimmer in voice produce mixed results. Adding other fractal dimension, the rate of recognition of the tremor segments in voice improves, but it still low (Teodorescu H.N. et al, 2005). This section of the database is useful in medical and phonological researches. Also for medical education use, the site comprises a gnathosonic and gnathophonic corpus. It offers opportunities for diagnosis and treating of speech by hearing the correct pronunciation of the words. In figure 2, we exemplify a gnathophonic (a) and gnathosonic (b) recording sounds (of the speaker 19743m). In figure 2(a), we exemplified recordings of the words: “vata”, “fata”, “var”.

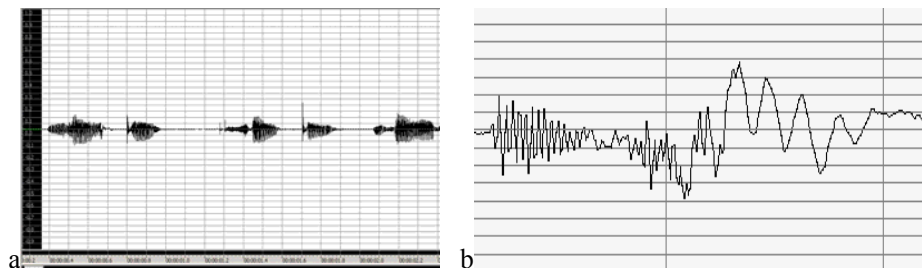


Figure 2. Gnathophonic (a) and gnathosonic (b) recording with details, tool GoldWave™. By analyzing such recordings available at SRoL, students can learn how to differentiate the normal and pathological states

Figure 2(b) illustrates a double occlusive sound. These two occlusive sounds are separated in time and denote two contact points. This is a type of occlusive sound which can produce in time a deficiency in the mandibular movement and erosion of teeth. The educators and students can use many the statistical studies regarding the pathological sounds in the Romanian language and recordings of persons with different pathologies (see section: Gnathosonic and Gnathophonic Archive) that the site includes.

Emotions rendering in voice and emotion analysis is increasingly addressed in recent years, including for medical and psychiatric diagnosis and treatment (Lundberg I., 1995), (Olofsson Å., 1992). The recordings from the emotional database at Max-Planck-Institute of Cognitive Neuroscience are made by a female fluent speaker; they made an electroencephalogram (EEG); the validation commission has twenty persons; they didn't offer information about the listeners; they judged the semantic content and the prosodic feature on five-point scale; the goal was to relate the emotions and to recognize a location in the human brain (Polzin, T.S., Waibel, A.H., 1998). We have addressed the topic in SRoL. The SRoL database contains feminine and masculine emotional voices; the speakers are aged between 25-35 years and they have no manifested pathologies. We analyzed only the audio voice signal. We did not make analyses like EEG, EMG, electroglottogram, etc., as those contained in other databases, like the Magdeburger Prosodie Korpus (Wendt B., Scheich H., 2002).

Every recording from the SRoL database is accompanied by the speaker profile and by the questionnaire concerning vocal pathology and objective factors for every speaker (Feraru, M., Teodorescu, H.N., 2008). The speaker's profile offers linguistic, ethnic, medical, educational, professional information about the speaker. The questionnaire presents details regarding the health state of the speaker (http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/ro/protocol_nou.htm).

4 Applications in teaching the voice signal technology classes

Signal technology classes are taught around the world. For examples, the Center for Spoken Language Understanding (CSLU) offers available language database from speech area and hearing science. These resources are important for analyzing the speech, for diagnosing and treating speech and language problems, for training students and so on. The tools and the corpora are distributed to over 2000 sites in 65 countries (Cole R.A., 1999). In education these tools help students learn about speech, learn a new language, learn through interactive media systems, or to become accustomed to hearing the normal and abnormal voice signal.

The SRoL team developed instruments for vocal signal processing regarding the extraction of patterns from this signal, and the computing of the fundamental frequency trace, respectively the traces of formants F1, F2, F3. The site offers, beside executables programs, descriptions for each of these tools. Those descriptions are intended for a "medium user", offering elementary explications and relevant references for a deeper understanding (Teodorescu H.N. et al., 2007), (Cristea D. et al, 2004).

The second author currently uses the SRoL corpus in teaching and laboratory activities in the class "Speech Technology" given for the master degree in "Computational

Linguistics” at the Faculty of Computer Science, “Al.I. Cuza” University of Iași. Details on the use in Voice Technology classes of some topics from SRoL are described in (Cristea D. et al, 2004). At the international EUROLAN 2007 summer school, the second author used the SRoL site to present “Traces of emotion, intentions and meaning in spoken Romanian” (<http://eurolan.info.uaic.ro/html/profs/HNTeodorescu.html>). The second author taught the specific methodology aspects, results obtained on the characterization of emotions in speech, possibilities of recognition of emotions and intentions in speech, and the relationship between specific meanings and the prosody in specific constructions in the Romanian language. The lesson exemplified applications of analysis of the speech emotional prosody to social, psycho-social, educational and psycho-medical topics.

5 Discussion and conclusions

Our team has a long standing experience with using novel technologies in teaching, hosting for three decades (Teodorescu H.N., Sofron E., 1987), (De Coulon et al., 1996), (Teodorescu H.N., 1998), (Teodorescu H.N. et al, 2000a,b), (Teodorescu H.N., 2001). We applied that experience to the SRoL e-teaching and e-learning resource.

The SRoL resource is a vast annotated corpus of speech files complemented by tutorials, papers and additional files, moreover with tools for speech processing. If used by an experimented student or teacher, it may become a powerful tool for instruction and learning the Romanian language pronunciation, speech technology, and voice pathology and re-education. The SRoL sound voice resource is useful in many domains, including phonology, applied computer science, and medicine. Students and researchers have the opportunity to have a freely accessible site, for learning the pronunciation of Romanian language, for making comparative study between Romanian and other language, for development of synthetic systems, for other linguistic, phonetic, socio-linguistic or medicine applications.

This database is structured corresponding to precise criteria, documented and annotated according to a well defined methodology. The site has more than 1500 recordings of syllable, word, sentence with different tonalities and pronounced with different emotional states. The database contains recordings of professional and normal voice, from the Nord – East region of Romania, without dialectal accent.

The SRoL resources have been recognized by several bodies, beyond the scientific publications that included our papers on SRoL. CLARIN European Network of Language Resources accepted SRoL as a member; ORDA (the Romanian Office for Authorship Rights) registered the original recordings, and the SRoL received a gold medal and media attention at the INVENTICA 2009 fair for inventions and creativity. Also, the website of Ambassade de France in Romania briefly described in its Bulletin the SRoL site and its use in education (<http://www.bulletins-electroniques.com/actualites/58811.htm>). Technical University „Gheorghe Asachi” of Iasi, Faculty of Electronics, Telecommunications and Information Technology) intends to use SRoL in helping foreign students enrolled at this university learn Romanian pronunciation.

We hope the SRoL resources will be used in all the universities in Romania by foreign students who learn the Romanian language, moreover in other academic media and as an online tool by foreign students and teachers. We welcome any request for help and educational advice from all those who wish to use our SRoL language-related web resources in virtual e-teaching and learning.

Acknowledgments. Research partly performed for the Romanian Academy “priority research” theme “Cognitive Systems” and to CEEEX grant nr 46/2005. We thank those who contributed to SRoL, primarily D. Trandabăț, M. Zbancioc, R. Luca, and L. Pistol.

REFERENCES

- Wendt, B. and Scheich, H. (2002): *The Magdeburger Prosodie-Korpus*. Proc. Speech Prosody Conf. Aix-en-Provence, France, pp. 699-701.
- Polzin, T.S. and Waibel, A.H. (1998): *Detecting emotions in speech*, Conference on Cooperative Multimodal Communication - Proc. CMC., Tilburg, The Netherlands.
- Cole R.A. (1999): *Tools for Research and Education in Speech Science*, Proc. Int. Conf. for Physics Students, www.cslu.ogi.edu/toolkit/pubs/pdf/cole_ICPS_99.pdf.
- Wise B.W. and Olson R.K. (1994): *Computer Speech and the Remediation of Reading and Spelling Problems*, J. Special Education Technology, vol. 12, nr. 3, pp. 207-220.
- Olofsson Å. (1992): *Synthetic speech and computer aided reading for reading disabled children*, Journal: Reading and Writing, vol. 4, nr. 2, pp. 165-178, ISSN: 09224777, (<http://www.springerlink.com/content/j521536n135x2864/>).
- Lundberg I. (1995): *The Computer as a Tool of Remediation in the Education of Students with Reading Disabilities: A Theory-Based Approach*, Learning Disability Quarterly, vol. 18, nr. 2, Technology for Persons with Learning Disabilities (1995), pp. 89-99, <http://www.jstor.org/pss/1511197>.
- Cameron K. (1999): *Computer Assisted Language Learning (CALL): Media, Design, and Applications*, Taylor & Francis, ISBN: 902651543X, http://www.google.com/books?id=dO_sNQIWhrsC&printsec=frontcover&dq=related:ISBN0940753030&hl=ro&source=gbs_similarbooks_s&cad=1
- Warschaw M. (1997): *Computer-Mediated Collaborative Learning: Theory and Practice*, The Modern Language Journal, Vol. 81, No. 4, Special Issue: Interaction, Collaboration, and Cooperation: Learning Languages and Preparing Language Teachers (Winter, 1997), pp. 470-481, <http://www.jstor.org/pss/328890>
- Solomon C. (1988): *Computer Environments for Children –A Reflection of Theories of Learning and Education*, www.google.com/books?id=EonPZ9A81kkC&printsec=frontcover&hl=ro&source=gbs_v2_summary_r&cad=0
- Feraru M. and Teodorescu H.N. (2008): *The Emotional Speech Section of the Romanian Spoken Language Archive*, Proc. 5th European Conf. on Intelligent Systems and Technologies, Iași, România, ISBN 978973730497.
- Cristea, D., Teodorescu, H.N., Tufiş, D.I. (2004): *Student Projects in Language and Speech Processing*, Workshop on Language Resources: Integration and Development in E-learning and in Teaching Computational Linguistics, pp. 17-22, 4th Conf. on Language Resources and Evaluation, Lisbon, Portugal - <http://nats-www.informatik.uni-hamburg.de/view/Main/AcceptedPapers>.
- Teodorescu, H.N., Tandabăț, D., Feraru, M., Zbancioc, M., Luca, R. (2007): *A corpus of the sounds in the Romanian spoken language for language-related education*. Ch. 6, pp. 73-89. In: C.P. Pascual (Ed.), “Revisiting Language Learning Resources”, Cambridge Scholars Pub. (CSP), UK, ISBN 1847181562; pp. 73-89.
- Teodorescu, H.N., Dobrea, D.M., Forte, E., Wentland-Forte, M. (2000a): *A High Sensitivity Sensor for Proximity Measurements and Its Use in Virtual Reality Applications*, Proceedings of the European Conference of Intelligent Technologies, Iași, România, ISBN 973-95156-1-4.
- De Coulon, F., Forte, E., Mlynek, D., Teodorescu, H.N., Suceveanu, Șt. (1996): *Subject State Analysis by Computer in CAE*, Proc. Int. Conf. on Intelligent Technologies in Human-Related Sciences, Leon, Spain. vol. 2, pp. 243-250.

- Teodorescu, H.N., Ganea, R., Feraru, M., Burlui, A. (2005): *Assesment of Voice quality based on nonlinear dynamic analysis*, Proceedings of The 15th Int. Conf. on Control Syst. & Computer Sci., Bucharest, România, pp. 536-542, ISBN 9738449898.
- Teodorescu H.N. (1998): *Computer semiotics: understanding meanings and parallel languages* (Refereed invited paper). In: T. Yamakawa, G. Matsumoto (Eds.): Proc. Int. Conf. IIZUKA'98, World Scientific Publ., pp. 279-283.
- Teodorescu, H.N., Kandel, A., Paschall, B. (2000b): *Teaching Modern Chapters in Automata Theory and Formal Languages*, Symp. 21 Century Teaching Technologies, March 20, Univ. South Florida, Tampa, USA (abstract in booklet of the Symposium).
- Zbancioc M. (2006): *Tools for the Archive of the Romanian Language Sounds Project*, 4th European Conf. on Intelligent Systems and Technologies, Iași, România, ISBN 973-730-265-6.
- Teodorescu, H.N., Zbancioc, M., Mihăilescu, E. (2006): *Speech Technology and Bio-Medical Engineering Teaching Based on the Web-A new Tool and Case Study*, International Conference on Interactive Computed Aided Learning, Sept. 27-29, Villach, Austria.
- Teodorescu, H.N, Feraru, M., Trandabăț, D. (2007b): *Studies on the Prosody of the Romanian Language: The Emotional Prosody and the Prosody of Double-Subject Sentences*, C., Teodorescu, H-N (Eds.) Advances in Spoken Language Technology, The Publishing House of the Romanian Academy, Bucharest, România, ISBN 978-973-27-1516-1, pp. 171-182.
- Teodorescu, H.N. and Feraru, M. (2007): *Micro-corpora de sunete gnatosonice și gnatofonice*, Pistol, Cristea, Tufiș (eds.) Resurse lingvistice și instrumente pentru prelucrarea limbii române. Editura Universității "Al. I. Cuza" Iași, pp. 21-30, ISBN 978-973-703-297-3.
- Teodorescu, H.N. and Feraru, M. (2008): *Classification in Gnathophonics – Preliminary Results*, The Second Symposium on Electrical and Electronics Engineering, Galați University Press, pp 525-530, ISBN 1842-8046.
- Teodorescu H.N. (2001): *Interrelationships Communication Semiotics*, in vol. "What Should be Computed to Understand and Model Brain Function: From Robotics, Soft Computing, Biology and Neuroscience to Cognitive Philosophy", (Ed.) Tadashi Kitamura. World Scientific, ISBN 9810245181, pp. 115-148.
- Teodorescu, H.N. and Sofron, E. (1987): *Demonstrating dislocations evolution*, Int. Journal Applied Engineering Education. vol. 3, nr. 2, pp. 189 -192.

Note. Due to the character of this article and to space limits, more references from the literature could not be included, as needed by the topic of the paper.