



South Asia Language Resource Center

SALRC GRANT LETTER OF INTENT 2005-2006

Proposal title	A Web-based Multimedia Pedagogical Dictionary of Telugu		
Language(s)	Telugu	Have you previously received SALRC funding for an unrelated project?	No
Project period (Earliest start date is 8/15/2005; project must be completed by 8/14/2006)	August 15, 2005 to August 14, 2006	Requested funding amount	\$ 19,000
Is this proposal an extension or expansion of a previously SALRC funded project?	No		
Faculty PI	Peri Bhaskararao		
Department	Linguistics	Office telephone	FAX
		+81-42-330-5607	+81-42-330-5607
Institution	Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies Asahi-cho 3-11-1, Fuchu-shi TOKYO 183-8534, Japan	E-Mail	peri@yhb.att.ne.jp bhaskar@aa.tufs.ac.jp bhperi@gmail.com
Faculty co-PIs (if applicable) and institutions			

Please respond to the following points (total 1000 words or less).

- Summarize the proposed work to be funded.
- How does this proposal fit SALRC's goals and how does it uniquely satisfy current language needs?



South Asia Language Resource Center

A Web-based Multimedia Pedagogical Dictionary of Telugu

PROPOSED WORK

Section A: About the Total Project

1. What is proposed (in brief):

We propose to construct a dictionary of Telugu that will be useful in learning and teaching the language. It will be a computer-based dictionary with text, audio and video files integrated into a coherent system. It will be primarily a web-based system that could be accessed over the internet and also could be supplied on a suitable media such as DVDs.

2. Details of the Proposal:

2.1. Contents of the Final Output: The result of this project will contain around 6000 records. The whole output will be structured for seamless export into a web document. Standard Generalized Markup Language (SGML) tags will be used for demarcating the various fields within a lexical record. Also, marking up will be done for efficient cross-linking of relevant records or relevant fields across different records.

Each record will have three components: Text, sound and picture (optional).

2.1.1. Text: The structure of the text part of a record is the following: Head-word; grammatical category information; meaning(s); illustrative phrase or sentences, where necessary. All portions in Telugu will be in UTF-8 Telugu encoding. The grammatical category information field will also contain information about oblique forms of nouns (where necessary) and conjugational classes of all verbs (for efficient inflection by the user).

2.1.2. Audio and Video: The structure of the audio portions of a record is: (a) audio file containing the pronunciation of the head-word, (b) pronunciations of an illustrative phrase or sentence (in two tempos: one utterance reflecting a slow word-by-word pronunciation and another utterance of the same item in a regular tempo reflecting the various morphophonemic changes that take place in natural pronunciation). Some of the words and phrases can be effectively illustrated by means of pictures – still or motion pictures. A digital still picture will be created where necessary. Some specific activities or movements require a video clip for illustration. Proper video files will be created for this purpose.

3. Need for this work:

There are very few published pedagogical materials of high quality available for non-Telugus, as demonstrated earlier by Bhaskararao (1974) and confirmed by a recent review of titles listed in WorldCat. Although we have conventional printed dictionaries like Gwynn (1991), now out of print, the proposed dictionary is unique and is essential. Thanks to the advances in multimedia computing, we can now seamlessly combine all the three essential components of language learning and teaching viz., textual, audio and video. Such a product brings the learner as close as possible to a native speaker of the language. With the aural materials available to the learner at the click of a button, one can develop sufficient oral skills of the language. Pronunciation of words in isolation and in context (in illustrative phrases or sentences) can be heard from a variety of speakers. A compact dictionary of this size will serve



South Asia Language Resource Center

the needs of the beginning and intermediate students of the language. The visual component of the dictionary will introduce the words and phrases in the proper cultural context where necessary.

This is a unique work for this language – there is no comprehensive work of this type available so far for Telugu. Thus there is no duplication of work involved. On the other hand, all the existing supporting materials will be properly taken into consideration while building up this work, including Prof. Velcheru Narayana Rao manuscript course book. The proposed dictionary will encompass all the vocabulary used and discussed in that course book.

4. Collaboration and Implementation:

This work will be done in collaboration with The Digital Dictionaries of South Asia of University of Chicago (DDSA). Collaborative discussions will be conducted at various stages of the work. The PI will prepare the base mark-up material for the dictionary (including the text and multimedia materials). DDSA will do the necessary browser-implementation. All the necessary technical specifications prescribed by the Networking Services & Information Technologies (NSIT) of the University of Chicago will be followed. The final output will be made available on the internet from minimum of two sites viz., University of Chicago and ILCAA, Tokyo University of Foreign Studies.

5. Post-Implementation By-products:

In addition to resulting in a pedagogically useful and essential multimedia dictionary, we consider that this project will facilitate establishment of standards that could be recommended as a frame of reference for future work on pedagogical dictionaries in multimedia context. The framework that will be laid down in this work could be successfully used as a template for future similar works.



South Asia Language Resource Center

6. Logistics of the Work:

6.1. Collection, Validation, and Processing of Data:

The major components include: 6.1.1. Text processing; 6.1.2. Text validation; 6.1.3. Audio-Video Data Collection; and 6.1.4. Audio-Video Data Processing. A detailed statement will be provided on each element of logistics and the related standards when a final proposal is submitted.

6.2. Presentation:

DDSA will prepare the material for presentation on the Web. Since the output of the processed data should match the input for the work by DDSA, all the necessary steps and specifications will be laid down beforehand.

Section B: Division of the Project into Parts

The total work will require about 20 months of time span. Hence, it is proposed to divide the work into two budgetary years. However, the output of the first year also could be seen as a functional module of the total work.

Part 1 of the Work

During the current budget year (August 2005-August 2006), the following modules of the work will be completed: Text-processing and Text-validation. The output of this module forms the base for the total work and requires more concentration. In addition to these two modules, we also will prepare a good number of samples of audio and video materials. Thus, the deliverables at the end of August 2006 will be the following: 1. A fully processed text, 2. Sample audio and video data processed and linked to the relevant portions of the text. DDSA will implement this part on the web. This will also give us an opportunity to make any fine tuning to the work that would feed into the next part.

Part 2 of the Work

Upon successful completion of Part 1, a request will be made for funding the second part. This part will contain the full portions of materials under 6.1.3 and 6.1.4. DDSA will then implement the full system on the web.

Reference:

- Bhaskararao, Peri. 1976. "Telugu courses for non-Telugus." *Telugu vani: monthly journal of the International Telugu Institute*, 1, 37-44.
- Gwynn, J. P. L. 1991. *A Telugu-English dictionary*. Delhi; New York: Oxford University Press. Web presentation by Digital Dictionaries of South Asia at <http://dsal.uchicago.edu/dictionaries/gwynn/> (accessed May 5, 2005)