

HQSYN16 - Task #4214

Task # 3679 (New): RA3d - Revision of positional parameters and weighting
Task # 4159 (New): Revision of continuous positional parameters

Redefine cost computation scheme in a language independent way

02.06.2017 12:02 - Tihelka Dan

Status:	New	Start date:	02.06.2017
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:			
Description			
Although the experiments described on positional features experiments display significant improvement of the last syllable placement, they are still related <i>only</i> to the position in the last syllable, which was designed with Czech in mind. The aim of this task is to design a new, language independent scheme which could be used for all ARTIC voices and languages. The key idea is to define a set of <i>significant positions</i> in a prosodic word (or any other rhythm unit). The position cost is then related to those significant positions. They may be stress or a last syllable nucleus in Czech, but any other feature in other languages. The position may also vary for individual prosodic words (e.g. where stress moves). The proposed scheme is as follows: • each candidate unit defines its relative position within the prosodic word $p(u)$ • each target unit defines its relative position within the prosodic word $p(t)$ • there is set of n significant point positions for the given prosodic word in target $s(t,1), s(t,2), \dots, s(t,n)$, each assigned with a weight $w(1), w(2), \dots, w(n)$ • also, each unit has its relation to the significant points in their corresponding prosodic word $s(u,1), s(u,2), \dots, s(u,n)$ the cost for i-th significant points is the given by a difference in distances from the point: • $vt(i) = \text{abs}(p(t) - s(t,i))$ • $vc(i) = \text{abs}(p(u) - s(u,i))$ • $\text{cost}(i) = \text{abs}(vt(i) - vc(i)) / \min(vt(i), vc(i)) + 1$ • and the total position cost is the sum through all the i significant points <i>Note:</i> thinking about it, the position features do not have to be shared across all the languages, i.e. each language/voice can have its own computation scheme. It just requires a more complex code handling the individual cases.			

History

#1 - 02.06.2017 12:04 - Tihelka Dan
- Description updated