

MAPEAMENTOS E TRANSCRIÇÕES FONÉTICA PARA O PORTUGUÊS BRASILEIRO

Verônica Feldkircher (BIC-UCS), André Gustavo Adami (orientador) - vfeldkir@ucs.br

O processamento de voz é um desafio a ser enfrentado pela computação, engenharia e lingüística no desenvolvimento de máquinas capazes de interagir de forma natural com os seres humanos. O projeto de pesquisa Bravoz surgiu da falta de recursos (tais como dicionário fonético, modelo de linguagem, modelo acústico, tabelas de conversão entre alfabetos fonéticos, alinhamento ortográficos e fonéticos) para sistemas de reconhecimento automático de fala contínua de grandes vocabulários (Large Vocabulary Continuous Speech Recognition – LVCSR) para o Português Brasileiro (PB) e da necessidade de um sistema LVCSR de referência com suporte a grandes vocabulários. Uma tabela de conversão de alfabetos fonéticos foi elaborada a partir dos símbolos fonéticos do IPA para o PB buscando correspondentes na tabela de símbolos fonéticos do Worldbet. Estes símbolos foram utilizados para realizar a transcrição e alinhamento fonético do corpus Spoltech. O processo de transcrição produz dois tipos de arquivos: um com indicação do conteúdo falado (transcrição ortográfica) e outro com a informação temporal dos fonemas (transcrição e alinhamento fonético), isto é, o símbolo fonético e o intervalo de tempo onde o fonema inicia e termina. Através desses dados um dicionário fonético é criado para que possa prover uma ou mais pronúncias para cada palavra do vocabulário e um modelo de linguagem é desenvolvido a partir do tratamento de bases de textos do PB para que reflita o conhecimento acerca da língua. O primeiro resultado foi a definição de um conjunto fonético para o PB e correspondência com o Worldbet, onde foi necessário criar novos símbolos que servissem às necessidades do PB anterior ao início do trabalho de transcrição e alinhamento fonético. Utilizando o novo conjunto fonético, foi realizada a revisão do alinhamento temporal fonético de 15% do corpus Spoltech (8119 arquivos de voz). A revisão elimina problemas de inconsistências encontrados em diferentes partes e permitirá a utilização eficaz da mesma para o desenvolvimento de sistemas LVCSR. Os recursos produzidos neste projeto facilitarão tanto o desenvolvimento de novos sistemas de reconhecimento automático de voz para o PB quanto o desenvolvimento de novas técnicas que são necessárias para o funcionamento de tais sistemas, servindo como base de dados e sistema de referência.

Palavras-chave: transição fonética, corpus de voz, fonética.

Apoio: UCS, CNPq.