

Telejornal para ler

Um sistema automático, acessível no teletexto, converte para legendas tudo o que é dito nas notícias

POR ALDA ROCHA

A Seleção Nacional de...» Mesmo depois de Portugal ter ficado pelo caminho no Euro 2008, há uma alta probabilidade de a palavra seguinte ser «futebol». E é isso mesmo que descobrirá o sistema automático de transcrição do *Telejornal*, desenvolvido pelo Laboratório de Sistemas de Língua Falada do INESC-ID Lisboa.

Parece simples, mas pôr uma máquina a reconhecer a fala significa um grande avanço tecnológico. Em casa, basta ir à página 885 do teletexto para poder acompanhar a legendagem, com um ligeiro atraso de segundos, tanto do que é dito pelo jornalista no estúdio como de todas as reportagens.

Mas, no laboratório, a realidade é um pouco diferente. O primeiro desafio do sistema é reconhecer o genérico, para começar a funcionar. «Basta-lhe um pequeno trecho da música para o reconhecer e fá-lo com bastante rigor», esclarece o investigador João Paulo Neto, 44 anos. É esta capacidade que permite suspender a transcrição durante o intervalo – não se pretende legendar a publicidade –, e parar antes do genérico final. Tudo sem intervenção humana.

Mas o maior desafio é transcrever as falas, o que exige o cruzamento de duas competências: identificar as palavras pelo som e pela probabilidade de ocorrerem. «Quando o sistema está bem treinado, temos uma taxa de reconhecimento fonético de 80%», explica o cientista. «Mas quando se junta o modelo de linguagem, conseguimos obter taxas muito mais elevadas.»

Cada pessoa tem uma forma única de falar, prolongando determinados sons e eliminando outros. Daí haver falhas. «Mesmo um humano, perante palavras isoladas, comete erros. Muitas vezes, precisa de as deduzir pelo contexto», explica a investigadora Isabel Trancoso, 52 anos.

O chamado modelo de linguagem consiste em conhecer a estrutura da língua, de modo a saber adivinhar uma palavra, por

exemplo, quando o ruído ambiente não permite percebê-la com clareza. Depois de «o primeiro-ministro José», será provável seguir-se «Sócrates». «Trata-se de uma gramática estatística», explica Isabel Trancoso. Para o sistema se tornar fiável foi alimentado com milhões de frases de textos retirados de notícias da Internet. Mas o grande salto de desempenho foi alcançado com horas de *Telejornais* transcritos à mão, por serem muito mais próximos da fala a «traduzir» todos os dias.

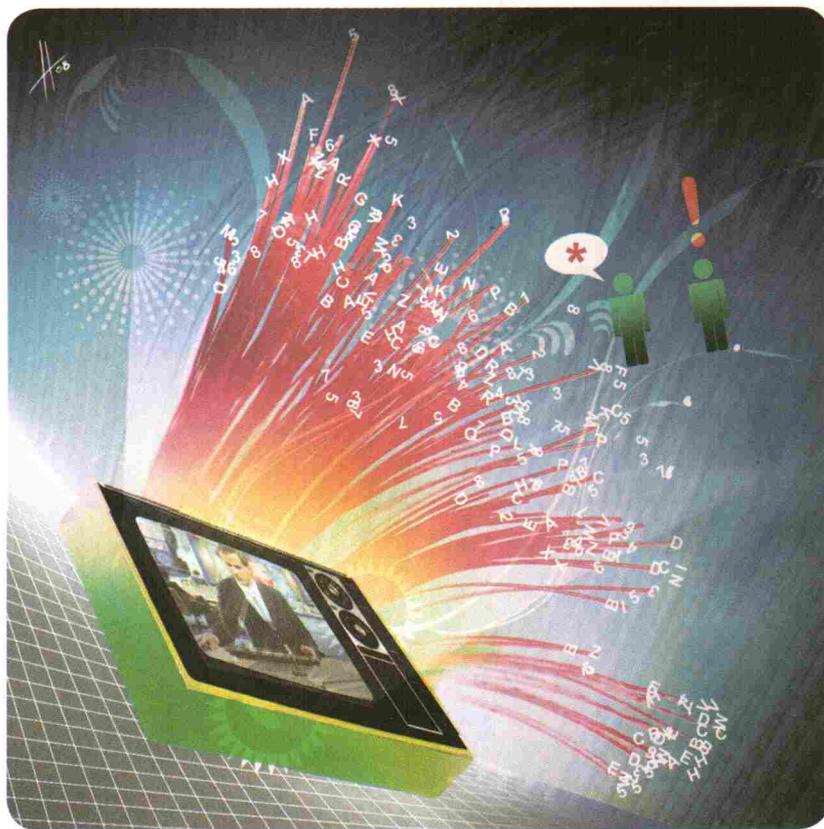
A DIFICULDADE DA FALA ESPONTÂNEA

O sonho de qualquer jornalista de ter um programa que converta entrevistas em texto parece muito mais perto, mas há ainda um grande obstáculo a ultrapassar.

«Sempre que existe fala espontânea, as taxas de erro são muito elevadas», afirma João Paulo Neto. No caso do *Telejornal*, as frases do *pivot* e a voz *off* das reportagens têm um reconhecimento muito bom, com 5% a 10% de erro. Mas, nas entrevistas de rua, há uma degradação. «Não só pelas condições sonoras – existe muito ruído ambiente a condicionar a recepção –, mas também pela qualidade da linguagem. Na fala espontânea, sucedem-se as hesitações, repetições, voltamos atrás...», esclarece Isabel Trancoso. E faltam textos escritos com o colorido típico de uma conversa imprevista, para ensinar o sistema. Nessas situações, a fim de evitar grandes

disparates, tem a «sensatez» de não legendar por não confiar no que está a «ouvir».

Não foi o que sucedeu nos primeiros testes, para risota dos cientistas. Logo no arranque do noticiário, José Alberto Carvalho deseja «Boa Morte». O sistema tratou de escolher o mais provável, já que «boa noite» não consta das notícias escritas. Uma vez mais, o universo do futebol ditava as regras – o computador escolheu o nome de um jogador. ▀



↳ Mesmo um humano, perante palavras isoladas, comete erros. Muitas vezes, precisa de as deduzir do contexto'

Isabel Trancoso, investigadora do Laboratório de Sistemas de Língua Falada do INESC-ID