

عنوان مقاله:

به کارگیری اطلاعات زبانی در یک سیستم بازشناسی گفتار پیوسته فارسی

محل انتشار:

دوفصلنامه زبان و زبان شناسی، دوره 6، شماره 11 (سال: 1389)

تعداد صفحات اصل مقاله: 26

نویسندگان:

محمد بحرانی - دانشگاه صنعتی شریف (آزمایشگاه پردازش گفتار)

حسین صامتی - دانشگاه صنعتی شریف (آزمایشگاه پردازش گفتار)

خلاصه مقاله:

در این مقاله یک سامانه بازشناسی گفتار پیوسته برای زبان فارسی معرفی می شود و نقش مدل آوایی و مدل زبانی در آن مورد بررسی قرار می گیرد. مدل های آوایی با روش های مستقل از بافت و وابسته به بافت در این سامانه به کار رفته و نتایج به کارگیری آن ها ارائه شده است. مدل زبانی سه کلمه ای نیز با روش های مبتنی بر کلمه، مبتنی بر مقوله نحوی و مبتنی بر طبقه، با استفاده از پیکره متنی زبان فارسی استخراج و در سامانه بازشناسی به کار گرفته شده است. همچنین مدل زبانی دستوری مبتنی بر دستور ساخت- گروهی تعمیم یافته در این سامانه پیاده سازی شده و نیز در ترکیب با مدل زبانی آماری به کار رفته است. نتایج حاصل نشان می دهد که مدل آوایی وابسته به بافت، مطابق انتظار، بهترین عملکرد را دارد. همچنین مدل زبانی سه کلمه ای مبتنی بر کلمه، نسبت به سایر روش های استخراج مدل زبانی آماری برتری دارد. در ضمن ترکیب مدل زبانی دستوری با مدل زبانی آماری منجر به بهبود نتایج بازشناسی می شود. سامانه بازشناسی گفتار معرفی شده در این مقاله، اولین سامانه بازشناسی برای گفتار پیوسته فارسی بوده و با پشتوانه فعالیت های تحقیقاتی متعددی که برای پیاده سازی آن انجام شده است، قابلیت استفاده به صورت کاربردی را یافته است.

کلمات کلیدی:

بازشناسی گفتار پیوسته، مدل سازی آوایی، مدل سازی زبانی، مدل زبانی آماری، مدل زبانی دستوری

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1581052>

