



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

تبدیل خودکار کلمه‌های لغت نامه دهخدا به قالب آوایی IPA

احمد یوسفان ، خدیجه جعفری ، محسن بیگلری

دانشگاه کاشان ؛ گروه کامپیوتر

yoosofan@kashanu.ac.ir

چکیده - آماده کردن آوای کلمه‌های فارسی یکی از پیش نیازهای بسیاری از کارهای پردازش گفتار فارسی و دیگر زمینه‌های پردازشی فارسی است. یک پایگاه داده‌ی آوایی کامل از کلمه‌های فارسی کمک شایانی به تبدیل متن فارسی به گفتار و شناسایی گفتار فارسی می‌نماید. در این مقاله روش نوینی برای ساخت خودکار پایگاه داده‌ی آوایی کلمه‌های فارسی به کمک کلمه‌های لغت نامه دهخدا پیشنهاد کرده‌ایم. در این روش کلمه‌های لغت نامه‌ی دهخدا و نشانه گذاری‌های ساده‌ی کلمه‌ها در این لغت نامه و برخی از ویژگی‌های آوایی کلمه‌های فارسی برای ساخت خودکار این پایگاه داده را به کار گرفته‌ایم.

کلید واژه - قالب آوایی ، IPA ، لغت نامه دهخدا ، پردازش گفتار

1- مقدمه

ای، او، آو، ای و اه) و سه حرکت دیگر به جای حرف مخصوص، علامت معینی دارد که باید در بالا یا زیر حرف رسم شود (،، ،) [۱] در این پروژه، برای تولید IPA هر کلمه از آوای ساده‌ای که در لغت‌نامه‌ی دهخدا برای هر کلمه در کنار آن نوشته شده است، کمک گرفته شده است. همچنین از برخی قانون‌های خواندن نوشته‌های فارسی که فارسی زبانان به کار می‌برند، نیز کمک گرفته شده است. از آنجا که زبان فارسی از کلمه‌های عربی و بیگانه‌ی دیگر نیز تشکیل شده است به دست آوردن قانون‌های تلفظ در آن مشکل است و کار را سخت‌تر می‌کند.

آوانویسی با الفبای IPA در واقع نشان دادن چگونگی خواندن کلمه است که در آن حروف لاتین به همراه علامت‌های ویژه‌ای به کار برده می‌شود. آوانویسی در شناخت دقیق تلفظ صحیح کلمه‌ها کاربرد دارد و امروزه در بیشتر فرهنگ‌های لغت این شیوه به کار برده می‌شود به طوری که امروزه یک برتری در ارزیابی فرهنگ لغت است. مهم‌ترین مشکل خط فارسی این است که برای همه‌ی حرکت‌ها حرف مخصوص ندارد، منظور حرکت‌هایی است که در لهجه عام یعنی لهجه‌ای که پایه آموزش فارسی زبانان بر آن نهاده شده است به کار می‌رود و آن‌ه حرکت است که شش‌تای آن حرف مخصوص دارد و جزء کلمه نوشته می‌شود (آ،

یکی از مسئله‌های دیگر در مورد این است که در لغت‌نامه‌ی دهخدا هم تلفظ قدیمی و هم تلفظ امروزی کلمه‌ها گنجانده شده است پس توجه به این امر هم ضروری است. الفبای IPA هر کلمه خروجی این برنامه است و بنابراین از نتیجه‌ی این پروژه می‌توان در جهت تبدیل متن به گفتار و شناسایی گفتار بهره برد.

2- لغت‌نامه دهخدا

لغت‌نامه دهخدا محصول مطالعات مداوم و کوشش‌های خستگی‌ناپذیر چهل و پنج ساله‌ی استاد دهخدا و گروهی از یاران اوست. وی در این مدت قریب سه میلیون «فیش» از روی متون معتبر استادان نظم و نثر دو زبان فارسی و عربی و لغت‌نامه‌های چاپی و خطی و کتب تاریخ و جغرافیا و علوم طب و هیئت و نجوم و ریاضی و حکمت و کلام و فقه و غیره فراهم آورده‌اند. این «فیش‌ها» اساس و پایه‌ی لغت‌نامه‌ی دهخدا را تشکیل می‌دهند.

شیوه ثبت لغات در کتاب لغت‌نامه دهخدا چنین است:

الف) لغت با حرف سیاه در اول سطر ثبت می‌شود و پس از آن نقطه‌ای گذارده می‌شود.

ب) در داخل [] حرکت حروف متحرک گذاشته می‌شود و از ضبط حروف ساکن صرف نظر می‌شود (مگر در صورت (التباس)). بدین طریق تلفظ هر کلمه به خوبی شناخته می‌شود. ماقبل «الف» (آ) اگر مفتوح باشد از ثبت آن خودداری می‌شود، همچنین ماقبل «ی» اگر مکسور باشد و ماقبل «و» اگر مضموم باشد. حرف مشدد در اینجا با تکرار آن نشان داده می‌شود: ماده [د د] ، و اگر متحرک باشد حرکت بر روی دومین حرف مکرر قرار می‌گیرد. «ی» که «ا» تلفظ می‌شود در اینجا به صورت «ا» نوشته می‌شود: یجیی [ی ی]. اگر «ها» در کلمه ملفوظ باشد آن «ها» را نیز بین القوسین قرار داده می‌شود مثلاً: اکراه [ه]، مه [م ه] و اگر تلفظ نشود نقل نمی‌شود مثلاً: عمله [ع م ل]، فاطمه [ط م].

ج) اصل کلمه که از چه زبانی است یاد می‌شود با علامت اختصاری و یا با نقل تمام حروف.

د) نوع کلمه از لحاظ دستوری با نشانه‌های اختصاری تعیین می‌شود، مثلاً (!) به جای اسم، (ص) به جای صفت ... (در مجلدات اول لغت نامه این امر رعایت نشده است). آنچه را که در این قسمت به نام «اتباع» خوانده می‌شود، منظور معنایی وسیع‌تر از آنچه عرب می‌گوید، می‌باشد. مقصود، هر دو کلمه‌ایست که غالباً با هم آیند اعم از آنکه هر دو با معنی باشند مانند «شک و ریب» و «یک و دو» یا یکی با معنی و دیگری بی معنی باشد مثل «یک و پهلو».

ه) سپس به شرح لغت پرداخته می‌شود با ذکر مآخذ.

و) آنگاه مترادفات کلمه نقل می‌شود.

ز) پس از آن شواهد نظم و نثر را (غالباً به ترتیب تاریخ) آورده می‌شود.

ح) معانی دیگر لغت با علامت || مشخص می‌شود.

ط) اعلام (اسماء خاص) مانند لغت نقل می‌شود و حتی‌المقدور مآخذ نیز در جای خود آورده می‌شود.

3- آواشناسی

آواشناسی مطالعه و توصیف علمی آواهای زبان و اصوات گفتاری است. آواشناسی ممکن است آواهای یک زبان بخصوص، مثلاً زبان فارسی باشد. و نیز موضوع آواشناسی ممکن است آواهای زبان بطور کلی باشد، این گونه بررسی‌های آوایی را آواشناسی همگانی می‌نامند.

3-1- رابطه‌ی آوا با خط

خط نیز مانند آوا به عنوان ماده‌ای که نظام‌های زبانی را در خود منعکس می‌کند و با زبان رابطه پیدا می‌کند. خط نشانه‌ی مادی، محسوس و نوشتاری برای نظام ذهنی زبان است. به این ترتیب، زبان یک پدیده ذهنی واحد با دو تظاهر مادی متفاوت است که یکی نوشتاری و دیگری گفتاری است. اگر چه کارکرد آوا و خط در زبان یکی است، اما نوع رابطه‌ای که بین خط و زبان موجود است با نوع رابطه‌ای که بین آوا و زبان وجود دارد بکلی متفاوت است؛ زیرا رابطه‌ی آوا و زبان مستقیم و بی‌واسطه است در صورتی که رابطه‌ی خط با زبان غیر مستقیم و با واسطه می‌باشد.



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

2-3- واج‌شناسی و فرق میان آواشناسی و

واج‌شناسی

مطالعه اصوات به عنوان کوچکترین واحدهای زبانی که از ترکیب آنها واژه‌ها و عبارتها ساخته می‌شوند، واج‌شناسی خوانده می‌شود. آواهای هر زبان را می‌توان از دو جهت مورد مطالعه قرار داد: نخست از جهت نحوه تولید و ویژگی‌های صوتی آوا مورد تحقیق واقع می‌شود. این‌گونه تحقیقات را که وظیفه اصلی آن بررسی و شناخت آوا به‌عنوان یک پدیده فیزیکی و مادی است، آواشناسی می‌نامند. دیگر اینکه آواهای موجود در آن زبان و نقش هر کدام از آنها در تمایز معنایی و قواعد ترکیب آنها معین شود؛ بررسی نظام صوتی یک زبان را واج‌شناسی می‌نامند. مطالعه آواشناسی برای واج‌شناسی مقدم است، زیرا پیش از شناخت آواهای یک زبان نمی‌توان به نظام موجود بین آنها و قواعد حاکم بر آن نظام دست یافت.

3-3- آوانویسی

برای اینکه بتوان نتیجه‌ی پژوهش‌های آواشناسی را به صورت نوشته عرضه کرد، به دستگاه الفبایی نیاز هست که آواشناس با استفاده از آن، آواهای زبان را طوری بنویسد که جز با یک تلفظ خوانده نشوند. این کار را نمی‌توان با کمک الفباهای موجود انجام داد، زیرا این نکته به صورت یک واقعیت عمومی پذیرفته شده است که همه الفباهای موجود عموماً دارای کمبودهایی هستند که با وجود آنها، از عهده‌ی ضبط دقیق زبان‌ها برنمی‌آیند.

4-3- الفبای آوانگار بین‌المللی (IPA)

به دلیل آن که آواشناسان بتوانند در سطح بین‌المللی از پژوهش‌های یکدیگر در این زمینه بهره ببرند، انجمن بین‌المللی آواشناسی (International Phonetic Association) (<http://www.langsci.ucl.ac.uk/ipa>) به این اندیشه افتادند که برای آواشناسی مجموعه‌ای از نشانه‌های قراردادی خاصی با ارزش‌های مشخص و تعریف شده وضع

کند. این فکر در سال 1888 میلادی جامه عمل پوشید، نتیجه آن دستگاه «الفبای آوانگار بین‌المللی» است که از آن سال تاکنون مورد استفاده‌ی آواشناسان بیشتر کشورهای جهان است و در صورت نیاز به روز می‌شود.

IPA بر اساس الفبای لاتین است و این برتری را دارد که زبانی آشنا و مانوس است اما همچنین حروف و نمادهای اضافه‌ای از منابع مختلف دیگر را در بر می‌گیرد. این حروف و نمادهای اضافه ضروری هستند، زیرا تنوع صداها در هر زبان بسیار زیادتر از تعداد حروف الفبای لاتین هستند. به کارگیری دنباله‌ای از نمادهای آوایی، «گفتار» را نمایش می‌دهد که معروف به «آوانویسی» است. IPA در مقاصد زیادی استفاده می‌شود؛ برای مثال، می‌تواند به عنوان راهی برای نمایش تلفظ در لغت‌نامه، ثبت یک زبان در زبان‌شناسی، به شکل درآوردن اساس سیستم نوشتاری یک زبان، یا تفسیر کردن صدا و نمایش‌های دیگر در تحلیل گفتار، استفاده شود. در همه‌ی این موارد داشتن یک مجموعه توافقی کلی از نمادها به منظور نمایش صداها به صورت غیر مبهم ضروری است، و IPA به انجام این نقش کمک می‌کند.

برای آشنایی بیشتر با آوانگاری و آوانویسی می‌توان به منابع گوناگونی مراجعه نمود برای نمونه [۲، ۳، ۴] چند نمونه از کتاب‌هایی در این زمینه هستند. در [۵] فهرست کامل آوانگاری همه‌ی زبان‌های زنده‌ی دنیا گذشته شده است که بر پایه‌ی مؤسسه‌ی بین‌المللی آوانگاری استاندارد آماده شده است.

4- گذری بر کارهای گذشته

مهم‌ترین کار در زمینه‌ی واج‌نویسی رایانه‌ای در زبان فارسی در پروژه‌ی «واژگان زبانی فارسی» در پژوهشکده پردازش هوشمند علائم انجام شده است. در این پروژه ۵۰ هزار کلمه‌ی فارسی واج‌نویسی شده‌اند. این مجموعه داده شاید بزرگترین مجموعه در دسترس تا کنون باشد. آن گونه که از

گزارش این پروژه به نظر می‌رسد واج نویسی به صورت دستی انجام شده است. یکی از مشکلات این پروژه واج نویسی به صورتی غیر از استاندارد IPA است و دشواری کامل کردن آن به روش دستی برای دیگر کلمه‌های فارسی است.

همچنین می‌توان به کارهای انجام شده در دبیرخانه شورای عالی اطلاع رسانی در کارگروه خط و زبان فارسی در محیط رایانه اشاره کرد. از آن جمله می‌توان به پروژه‌های «دادگان صوتی تصویری تلفظ واجهای فارسی» و «تهیه واژگان زایا» در سال ۸۳ اشاره کرد.

5- روند انجام کار

در آغاز برنامه‌ای به زبان php نوشته شد تا کلمه‌های درون پایگاه داده‌ی هوشمند علائم خوانده شود و واج نویسی آن به صورت خودکار به واج نویسی استاندارد IPA تبدیل شود. این کار به نسبت ساده انجام شد زیرا نگاشت تقریباً یک به یک میان واج نویسی غیر استاندارد پژوهشکده‌ی هوشمند علائم و IPA وجود داشت. به این ترتیب واج نویسی استاندارد نزدیک به ۵۰ هزار کلمه آماده شد. همچنین این داده‌ها در پایگاه داده‌ی MySQL گذاشته شدند تا در صورت نیاز به سادگی بتوانند در پروژه‌های دیگر نیز به کار گرفته شوند و امکان تبدیل آن به دیگر DBMSها نیز وجود داشته باشد.

در این گام به کمک کلمه‌های موجود در لغت نامه‌ی دهخدا که درون نزدیک به ۲۵۰ هزار پرونده‌ی html گذاشته شده است، واج نویسی به صورت خودکار برای همه‌ی کلمه‌های فارسی این لغت نامه انجام شد. در این مرحله در آغاز کلمه‌های درون پرونده‌ها html به همراه حرکت‌گذاری آنها و نوع کلمه و البته توضیح‌های هر کلمه به صورت خودکار به کمک برنامه‌ای که نوشته شد، استخراج شد و درون جدولی در MySQL گذاشته شد. جزییات انجام این کار و برنامه‌ی نوشته شده در [۶] گذاشته شده است.

5-1- فشرده‌ای از گام‌های انجام کار

در ادامه برنامه‌ای برای ساخت خودکار واج‌ها از روی این جدول آماده شد که در ادامه توضیح داده می‌شود. برای اینکه توضیح روشن شود نمونه‌ای از کلمه‌های درون این

جدول در زیر نوشته شده است.

لغت: پنبه دوزی

آوا: [پِمَبَ / ب]

نوع: حامص مرکب

شرح: طرز دوختن جامه‌ای...

ساخت خودکار واج‌ها در شش گام زیر انجام شد:

۱- یکم: تجزیه کلمه به حروف تشکیل دهنده آن

۲- تجزیه آوای کلمه به حروف و اعراب موجود در آن

۳- ترکیب کلمه و آوای به دست آمده از دو گام پیشین

۴- اعمال قوانین واج شناسی روی نتیجه گام سوم

۵- تبدیل حروف فارسی به حروف معادل آنها در الفبای IPA

۶- ساخت الگوی هجایی و تعیین جایگاه تکیه‌ی متن آوایی

موارد مهمی که در همه‌ی این گام‌ها دنبال می‌شود به قرار زیر است:

۱- تعیین اعراب همخوان‌ها به کمک آوای ساده‌ی هر کلمه

۲- تعیین آوای حرفی که با یک نشانه برای آواهای مختلف استفاده می‌شوند، مثل حرف "و" یا حرف "ی"

۳- به کار بستن قانون‌های واج شناسی که به موجب در کنار هم قرار گرفتن واج‌ها می‌بایست رعایت شوند، به عنوان مثال: اگر واکه بلند "ای" در کنار واکه بلند "ا" قرار گیرد، همخوان "ی" بین آن دو قرار می‌گیرد (مثل نیایش) و ...

برای خلاصه شدن روند کار و نیاز نبودن به توضیح فراوان روند انجام گام‌ها در نمونه‌های زیر نشان داده می‌شود.

5-2- گام یکم و دوم: تجزیه‌ی کلمه و آوا به

حروف تشکیل دهنده آنها

کلمه‌ی «باد بهاری» [دِبَ] « به صورت زیر تبدیل می‌شود



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

(یا به عبارت دیگر بدون حرکت می‌باشد) زیرا در فارسی دنباله‌ی واجی هر حرف به جز برای حروف تشدید دار حداکثر یک همخوان (C) و یک واکه (V) به صورت /CV/ می‌تواند باشد. البته در این گام همه‌ی قانون‌های استثنای این قانون نیز در نظر گرفته شده است که برای طولانی نشدن مقاله، توضیح داده نمی‌شوند.

ب	ا	د	ب	ه	ا	ر	ی
---	---	---	---	---	---	---	---

به همراه

د	ب
---	---

3-5- گام سوم: ترکیب کلمه و آوا

ب	ا	د	ب	ه	ا	ر	ی
---	---	---	---	---	---	---	---

در این گام برخی از قانون‌های فارسی به کار گرفته می‌شود و بر روی برخی از حرف‌ها حرکت گذاری انجام می‌شود. این کار به صورت حرف به حرف بر روی آرایه‌ی حروف کلمه که از گام پیش به دست آمده است انجام می‌شود. در ادامه چند نمونه قانون به کار رفته آورده می‌شود.

۲- اگر حرف در کلمه برابر با حرف در آوا نباشد آن‌گاه کار دیگری انجام نمی‌شود و حرف پس از آن پردازش می‌شود مگر در چند مورد ویژه که قانون‌های ویژه‌ای برای خود دارند و در اینجا به آنها نمی‌پردازیم.

البته مورد‌های ویژه‌ی بسیاری نیز در این گام در نظر گرفته شد که تعدادی از آنها به قرار زیر است.

- «ال» عربی در کلمه و تلفظ ویژه‌ی آن. مانند «علی الطلوع» با آوای [عَ لَطُ طْ].
- حرف‌های با چند تلفظ
- لبی شدن حرف «ن»
- نوشته نشدن همزه در لغت نامه دهخدا

۱- حرف در کلمه برابر با حرف در آوا باشد. باید بررسی‌های زیر برای این کار انجام شود تا دقت آن افزایش یابد

۱-۱- اگر حرف پس از حرف کنونی برابر با واکه‌های بلند «آ ، و ، ی» نباشد آن‌گاه با توجه به مورد‌های زیر احتمالاً این حرف همین آوا را دارد

۱-۱-۱- اگر حرف کنونی تشدید داشته باشد آن‌گاه باید حرکت تشدید برداشته می‌شود و حرف دو بار در آرایه گذاشته شود که حرکت یکمی ساکن است و حرکت دومین حرف فته است. برای نمونه «محمد» تبدیل می‌شود به [م ح م م د].

4-5- گام چهارم: اعمال قوانین واج شناسی

روی ترکیب کلمه و آوا

برخی از کارهایی که در این گام انجام می‌شوند به قرار زیر است.

۱- اگر در کلمه‌ای پس از واکه «ی» بلافاصله یکی از واکه های «ا ، ، ، ، ا» قرار داشته باشد در هنگام تلفظ، حرف همخوان میانجی «ی» قرار می‌گیرد. به عنوان مثال: کلمه «ریاضت» با آوای [ضْ].

۲- تشخیص ملفوظ یا غیر ملفوظ بودن حرف /ه/ در پایان کلمه.

۱-۲- اگر حرف دارای دو تلفظ باشد در لغت نامه به کمک / دو تلفظ متفاوت آن از هم جدا شده است و در این جا برای کل کلمه دو تلفظ (یا بیشتر بسته به تعداد / در این حرف و در حرف‌های دیگر) گذاشته می‌شود. برای نمونه «دریای شمال» به صورت [د ر ی ش / ش] و «دلیری نمودن» به صورت [د ن / ن / ن] نوشته شده است.

۱-۲- اگر حرف پس از حرف کنونی برابر با واکه‌های بلند «آ ، و ، ی» باشد آن‌گاه این حرف دارای حرکت ساکن می‌باشد

۱-۲- حرف /ه/ غیر ملفوظ است در صورتی که آخرین حرف

در آرایه باشد و هیچ کدام از واکه های بلند /، و/ قبل از آن قرار نداشته باشد. مثل لغت "خنده" با آوای [خَ دِ] ، "درمانده" با آوای [دَ دِ/ دِ]

۲-۲- در غیر این صورت /ه/ ملفوظ است، به عنوان مثال: لغت "درمانگاه" با آوای [دَ]، یا لغت "کوه" یا لغت "روبه" با آوای [بَ هَ].

5-5- گام پنجم: تبدیل حروف فارسی به حروف

معادل آنها در IPA

در این گام یک نگاشت از آرایه‌ی آوای به دست آمده به آرایه‌ی IPA انجام می‌شود. برای نمونه در گام چهارم برای کلمه‌ی «توحید» آرایه‌ی آوای زیر به دست می‌آید.

ت	و	ح	ی	د
---	---	---	---	---

پس از تبدیل به IPA به آرایه‌ی زیر تبدیل می‌شود.

t	æ	v	h	i	d
---	---	---	---	---	---

در کلمات دارای حروف چند تلفظی، مثل کلمه "شکوفه" با آوای [شِ / شِ / فِ / فِ / فِ] در این مرحله تبدیل به چهار آوا می‌شود.

5-6- الگوی هجایی در زبان فارسی و مکان

تکیه

الگوی هجایی در کلمه‌های فارسی به صورت (C)V(C)(C) است که با در نظر گرفتن همزه در آغاز کلمه ای که با مصوت آغاز می‌شود، الگوی هجایی به صورت CV(C)(C) است. برای نمونه کلمه‌ی توحید به صورت زیر است.

T	æ	v	h	I	d
c	v	c	c	v	c
	W		l	S	

بنابراین این کلمه به دو بخش بالا شکسته می‌شود که به کمک آن می‌توان تا اندازه‌ای مکان تکیه (S) را در کلمه شناسایی نمود.

6- نتیجه‌گیری

در این مقاله روش کار برنامه‌ای توضیح داده شد که به

کمک آن به صورت خودکار قالب آوایی همه‌ی کلمه‌های درون لغت نامه دهخدا آماده شد. این روش برای ساخت خودکار واج دارای برتری‌ها زیر است:

۱- این روش به خوبی می‌تواند با بهبود قانون‌های به کار رفته در آن کارآیی بسیار بهتری پیدا کند چون در این روش با کمک گرفتن از آوای ساده درون لغت نامه و به کارگیری قانون‌هایی برای انجام تبدیل، آوای استاندارد به دست می‌آید.

۲- آماده شدن آوای استاندارد IPA برای بیش از ۲۵۰ هزار کلمه بدون نیاز به ساخت دستی این آوا و با زمان اجرای مناسب

مشکل اصلی این روش وابستگی این برنامه به آوای ساده‌ی کلمه‌ها در لغت نامه دهخدا است. امکان وجود خطا و تنوع در نوشتن آوای ساده برای این برنامه مشکل آفرین است و به این دلیل شاید نتیجه‌ی مناسب برای برخی از کلمه‌ها به دست آید.

این برنامه بر روی ۲۷۱ هزار کلمه اجرا شد و IPA استاندارد آنها را به دست آورد. برای بررسی دقت کار این برنامه به صورت تصادفی ۳۰۰۰ کلمه برگزیده شد و به صورت دستی بررسی گردید و ۹۵ درصد این کلمه‌های برگزیده شده دارای IPA ساخته شده‌ی درست بودند. به این ترتیب می‌توان به خوبی این روش را به عنوان یک روش ساخت آوای استاندارد به کار برد و بهبود بخشید و برای دیگر لغت‌نامه‌های در دسترس نیز گسترش داد.

مراجع

- [1] بهمنیار، احمد. «مقدمه لغت نامه دهخدا بخش املای فارسی». تهران: دانشگاه تهران ۱۳۷۳
- [2] فرامکلین، ویکتوریا ؛ رابرت رانمن ، نینا هایمز. مترجمان: علی رحیمی، رامین رهنورد. «مقدمه‌ای بر زبان». تهران: جنگل ، جاودانه ۱۳۸۷
- [3] حق شناس، علی محمد. «آواشناسی(فونتیکی)». تهران: نشر آگاه ۱۳۸۶.
- [4] ثمره، یدالله؛ امید طیب زاده ، رضا نیلی‌پور. «آواشناسی زبان فارسی: آواها و ساخت آوایی هجا» تهران: مرکز نشر دانشگاهی ۱۳۸۵.
- [5] International Phonetic Association. Handbook of the International Phonetic Association : A Guide to the Use of the International Phonetic Alphabet. Cambridge University Press (1999).
- [6] جعفری، خدیجه. تبدیل خودکار لغات موجود در لغت‌نامه دهخدا به فرم آوایی. IPA. گزارش پروژه‌ی کارشناسی، زیر



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

نظر احمد یوسفان، دانشگاه کاشان، دانشکده‌ی مهندسی گروه
کامپیوتر ۱۳۸۸