

بازنمایی دانش دستوری زبان فارسی به کمک گرامر پیوندی

احمد عبدالله زاده بارفروش

آرمین سجادی

ahmad@aut.ac.ir

asajadi@aut.ac.ir

دانشکده کامپیوتر، دانشگاه صنعتی امیرکبیر

بازنمایی دستور زبان را می‌توان به‌عنوان مهمترین فاز در سیستم‌های پردازش زبان طبیعی به‌شمار آورد. گرامرهای به‌کار رفته در این سیستم‌ها عموماً گرامرهای مبتنی بر واحد یا ساختاری می‌باشند که می‌کوشند جمله را با تجزیه به واحدهای ریزتر تحلیل نمایند. دیدگاه‌های متفاوتی برای اینکار وجود دارد که گرامر پیوندی به‌عنوان یکی از آنها مبنا را بر ایجاد پیوندهایی بین کلمات قرار می‌دهد و به همین دلیل گرامر لغوی نامیده می‌شود. این گرامر برای زبانهای متعددی ارزیابی شده و بررسی نحوه عملکرد آنها برای زبان فارسی برای اولین بار هدف این مقاله است. در این تحقیق خواهیم کوشید با شروع از روابط اصلی و اساسی در زبان فارسی، توسعه کار را مبتنی بر این روابط قرار دهیم و در هر مرحله تعداد این روابط را افزایش دهیم. برای اطمینان از انسجام مجموعه، قوانین در هر مرحله به کمک پارسر گرامر پیوندی و روی یک بستر تست مناسب تست شده است.

کلمات کلیدی: پردازش زبان طبیعی، گرامرهای محاسباتی، گرامر پیوندی، زبان فارسی

۱ مقدمه

آنچه دانش را از مفاهیم مشابه دیگری چون داده و اطلاعات متمایز می‌سازد سطح انتزاع در آنها می‌باشد [عبدالله زاده ۱۳۸۳] و دانش دستوری از مهمترین دانشهایی است که در هر سیستم پردازش زبان طبیعی وجود دارد. تاکنون فرمالیسمهای متعددی برای بازنمایی دستور معرفی شده که مهمترین آنها قوانین بازنویسی^۱ زبانهای مستقل از متن می‌باشد که به‌طور مستقل در [Chomsky 1957] برای بازنمایی زبان طبیعی و در [Naur 1958] برای استفاده در پارس زبانهای مصنوعی بکار رفت و به BNF^۲ مشهور است. استفاده از این فرمالیسم دو دلیل عمده دارد (۱) به اندازه کافی توانا هستند تا زبان طبیعی را بازنمایی کنند (۲) به اندازه کافی محدود هستند تا به سادگی پارس شوند [عبدالله زاده ۱۳۸۳]. این فرمالیسم مبتنی بر تجزیه سلسله مراتبی جمله به واحدهای کوچکتر می‌باشد و به همین دلیل تحت عنوان تجزیه سازه‌ای بلافصل^۳ و پارس مشهور است که زبان‌شناسان موسوم به ساخت‌گرایی آمریکایی برای توصیف نحو از آن استفاده کردند [دبیرمقدم ۱۳۸۳]. این

1) Rewrite Rules 2) Backus-Naur Form 3) Immediate Constituent Analysis

نظریه علیرغم داشتن تنوعهای گوناگون همگی در این خصیصه مشترک هستند و بهمین دلیل چامسکی همه آنها را گرامرهای با ساخت گروهی نامیده است و همچنان بخش اعظمی از گرامرهای محاسباتی را تشکیل می‌دهند. از مهمترین گرامرهای معرفی شده بر این اساس گرامرهای با ساختار گروهی تعمیم یافته^۱ [GPSG] (Gazdar et al. 1987) و نوع متاخر و بسیار پیچیده‌تری از آن با عنوان گرامرهای ساختار گروهی هسته-بنیان^۲ (HPSG) [Pollard and Sag 1994] می‌باشد. از سال ۱۹۸۰ به بعد، بیشتر تلاشها به سمت بازنماییهای با واژگان پیچیده‌تر و گرامرهای کمتر هدایت شده است [Russell and Norvig 2003]. اولین دسته از این گرامرهای مبتنی بر لغت، گرامرهای نقشی لغوی^۳ [Kaplan and Bresnan 1982] می‌باشد. از دیگر گرامرهای مبتنی بر این مفهوم می‌توان به گرامرهای درخت افزایشی^۴ [Joshi 1985] (TAG) اشاره کرد. لازم به ذکر است که هیچ یک از این گرامرها کاملاً لغوی نیستند و هر دو از مفهوم واحد برای تحلیل استفاده می‌کنند. از جمله گرامرهایی که در آنها اثری از واحد نیست گرامرهای وابستگی [Meščuk 1988] و گرامر پیوندی است. تاکنون در سیستمهای پردازش زبان فارسی روشهای چون قوانین بازنویسی مستقل از متن [فیلی و قاسم ثانی ۱۳۸۳]، گرامرهای درخت-افزایشی [فیلی و قاسم ثانی ۱۳۸۲] و HPSGها [آیت ۱۳۸۰] بکار گرفته شده ولی استفاده از گرامرهای لغوی سابقه چندانی ندارد و ما در این تحقیق خواهیم کوشید با اضافه نمودن مجموعه پیوندهای مورد نیاز، گرامر پیوندی را برای این کار تطبیق و کارایی آنها را مورد آزمون قرار دهیم. ما در بخش ۲ این گرامرها را بصورت مختصری معرفی می‌کنیم و با ترمینولوژی آنها در بخش ۳ آشنا می‌شویم. در بخش ۴ با شروع از نقشهای اولیه و الگوهای ساده در هر مرحله سعی در پیچیده‌تر کردن این تعاریف می‌کنیم تا در نهایت یک دیکشنری مختصر برای پردازش الگوهای ساده زبان فارسی ارائه دهیم.

۲ معرفی گرامر پیوندی

[Radford 1981] مهمترین مزیت گرامرهای مبتنی بر واحد را سه مورد عنوان می‌کند: اول تحت پوشش قرار دادن نظم توزیعی^۵ زبان، دوم قابلیت بازگشتی بودن آنها و سوم مقید بودن آنها برای کنترل پیچیدگیهای زبان. اما این نظریه با انتقاداتی نیز همراه شد یکی از آنها [Hudson 1984] و دیگری [Meščuk 1988] بود که به جای تکیه بر واحدها، استفاده از نوعی وابستگی ساختاری بین اجزاء جمله را پیشنهاد کردند. مثلاً در این روش از یک گراف جهت‌دار مسطح بین کلمات استفاده کرد که تشکیل یک درخت را می‌دادند. بجز ریشه، همه کلمات دارای یک یال به سمت فقط و فقط یک کلمه دیگر داشتند و هیچ یالی به ریشه ختم نمیشد. گرامر پیوندی (LG) علاوه بر لغوی بودن هر سه شرط مطرح شده برای گرامرها را دارا هستند [Sleator and Temperley 1993]. به هر حال گرامر پیوندی را باید ادامه این تلاشها و راهی برای استفاده از خود کلمات به جای واحد دانست. LG عملاً بر این فرض استوار است که یک وابستگی ساختاری زیرین برای زبان طبیعی وجود دارد. یک نکته را نباید از نظر دور داشت و اینکه گرامر پیوندی خود یک فرمالیسم جدید برای بازنمایی زبانهای مستقل از متن می‌باشد [Sleator and Temperley 1991] اما برای مقایسه بین دو فرمالیسم مختلف نیز ملاکهای موجود است از جمله [Sleator and Temperley 1993].

1) Generalized Phrase Structure Grammar 2) Head driven Phrase Structure Grammar
3) Lexical Functional Grammars 4) Tree Adjoining Grammar 5) Distributional regularities

• سادگی در بازنمایی

• کارایی عملیات پارس

• مفید بودن خروجی پارسر برای پردازشهای بعدی

به این ترتیب می‌توان مشخصه‌های گرامر پیوندی را بصورت زیر بر شمرد

• ساده سازی گرامر

• امکان ارائه یک مدل احتمالاتی [Lafferty et al. 1992].

• برخلاف گرامرهای مبتنی بر ساختار گروهی، کلماتی که از لحاظ نحوی یا معنایی بهم مرتبط هستند، بعد از پارس به هم پیوند می‌شوند و این مسئله رعایت تطابق را ساده‌تر می‌کند. همانگونه که خواهیم دید این ویژگی مدیریت تطابق در زبان فارسی را که بسیار لغوی است ساده‌تر می‌کند.

• یک ویژگی دیگر گرامر پیوندیها اینست که علیرغم قائل نبودن به واحدها، اطلاعات مربوط به آنها بسادگی قابل استخراج می‌باشد.

گرامر پیوندی در سطح وسیعی برای زبان انگلیسی بکار رفته اند. تعریف کاملی از واژگان برای بازنمایی بخش وسیعی از زبان انگلیسی در [Sleator and Temperley 1991] آمده است. از گرامر پیوندی برای استخراج رویدادها [Hashara et al. 2003]، استخراج دانش از سیستمهای اطلاعاتی پزشکی [Diang et al 2003] و موارد زیاد دیگری استفاده شده است. در سطح زبانهای دیگر نیز مورد توجه قرار گرفته است، برای نمونه می‌توان به زبان فرانسوی [Aubin 2003] و چینی [Liu 2002] اشاره کرد. همچنین گرامر پیوندی احتمالاتی^۱ [Lafferty et al. 1992] نیز توانایی زیادی را در زمینه پردازش گفتار نشان داده است [Hunt. 1993] و ارزیابی عملکرد آنها برای زبان فارسی هدف این مقاله می‌باشد.

۳ نمادگذاری گرامرهای پیوندی

در این بخش ترمینولوژی گرامر پیوندی را در حدی که در این پژوهش مورد استفاده قرار گرفته بررسی می‌کنیم، برای جزئیات بیشتر می‌توان به [Sleator and Temperley 1991] و [Sleator and Temperley 1993] رجوع کرد.

پیوند پیوندها جفت کلمات را به هم مرتبط می‌کنند که به هر یک اسمی اختصاص داده می‌شود.
اتصالگر پیوندها از طریق اتصالگرهای کلمات به آنها متصل می‌شوند که از آن به عنوان ارضا کردن اتصالگر ذکر می‌شود. هر اتصالگر شامل یک نام پیوند به همراه + یا - می‌باشد، دو اتصالگر هنگامی به هم متصل می‌شوند که اسمی آنها با هم مطابق بوده و اتصالگر +(-) در سمت چپ (راست) پیوند قرار گیرد.

ملزومات پیوندی تعریف پیوندها برای یک اتصالگر به شکل فرمولهای خوش فرمی است که توسط ترمهای تابعی ساخته می‌شود که لیست آنها در جدول ۱ آمده است.

1) Probabilistic Link-Grammars

تطابق پیوندی در حالت کلی یک اتصال گر با یک یا بیشتر حرف بزرگ شروع می‌شود که نام آن را مشخص می‌کنند و با دنباله ای از حروف کوچک ادامه می‌یابد که زیرنویس^۱ نامیده می‌شود. قانون کلی تطابق بین دو اتصالگر را می‌توان به فرم زیر خلاصه کرد: دو اتصالگر $x_1 \dots x_m$ و $y_1 \dots y_n$ قابلیت اتصال دارند هرگاه

(الف) A با B مطابقت کند و

(ب) $x_1 \dots x_n$ با $y_1 \dots y_n$ مطابقت کند و $m = n$ یا $x_1 \dots x_m$ با $y_1 \dots y_m$ مطابقت کند و $m < n$

هر پیوند بندی معتبر باید قوانین زیر را رعایت کند.

مسطح بودن^۲ پیوندها هنگام ترسیم در بخش بالایی کلمات نباید همدیگر را قطع کنند.

همبندی^۳ پیوندها باید همه کلمات را به هم متصل کنند.

ارضای^۴ پیوندها بایستی ملزومات پیوندی همه کلمات را ارضا کنند.

این گرامرها به علت تکیه کامل بر لغات، گرامرهای لغوی^۵ نامیده می‌شوند. با توجه به گستردگی گرامر در سطح لغات هر قانون در اینجا یک تعریف^۶ نامیده می‌شود و به مجموعه این تعاریف در نهایت دیکشنری گفته می‌شود.

۴ استفاده از گرامر پیوندی برای بازنمایی قوانین فارسی

در این بخش سعی می‌کنیم الگوهای اساسی دستور زبان فارسی را به کمک گرامر پیوندی بازنمایی کنیم و قابلیت آنها را برای اینکار مورد سنجش قرار دهیم، تلاش ما بر این بوده است که توسعه کار قدم به قدم صورت گیرد تا معیاری جهت ارزیابی مقیاس پذیری گرامر پیوندی بدست دهد. مبتنی بودن بیشتر مطالعات سنتی دستور زبان فارسی بر رابطه‌ها از یکسو و نقش محوری و مرکزی رابطه‌ها در دستور زبان که اساس کار دستورهای رابطه

جدول ۱: عملگرهای گرامر پیوندی

عملگر	کارکرد
&	هر دو اتصالگر باید ارضا شوند.
or	یکی از اتصالگرها بایستی ارضا شود.
{exp}	عبارت اختیاری است.
@ (unary)	یک یا تعداد بیشتری از این اتصالگر می‌تواند ارضا شود.
<MAC_NAME> : exp	عبارت کوتاه <MAC_NAME> می‌تواند بجای عبارت طولانی تر exp بکار رود که به آن ماکرو گفته می‌شود.

1) Subscripts 2) Planarity 3) Connectivity 3) Connectivity 4) Satisfaction 5) Lexical Grammars 6) Definition

ای را تشکیل می‌دهد [دبیر مقدم ۱۳۸۲] از سوی دیگر، گرامر پیوندی را به عنوان فرمالیسمی که در آن رابطه‌ها به سادگی از نتیجه پارس قابل استخراج هستند، مناسب می‌نمایاند. در این مقاله ما نیز تکیه را بر این رابطه‌ها قرار داده ایم و با شروع از مجموعه اولیه {فاعل، فعل} در هر مرحله یک یا چند رابطه را به این مجموعه اضافه می‌کنیم. با توجه به اینکه هر رابطه عموماً با یک پیوند بین یک جفت کلمه تعیین می‌گردد می‌توان از یک نگاه هر مرحله را افزودن پیوندی به گرامر محسوب کرد. اگر فرض کنیم منظور از $X(\alpha, \beta)$ اینست که دو رابطه α و β از طریق پیوند X به هم متصل می‌شوند و A_{t+1} را مجموعه پیوندهای تعریف شده تاکنون قرار دهیم، هر مرحله می‌تواند به صورت $\{X(\alpha, \beta)\} \cup A_t = A_{t+1}$ نشان داده شود. منظور از الگو زوجی چون (B, f) است که در آن $B \subseteq A$ و f تابعی است که نحوه حضور هر عضو از B را در پارس جمله نشان می‌دهد. مبنای استخراج الگوهای اولیه [خانلری ۱۳۸۲] می‌باشد و در ادامه دیگر الگوها از [انوری و گیوی ۱۳۸۳] استخراج شده است. در بازنمایی دستگاه مطابقت و نیز گروه‌های فعلی از [باطنی ۱۳۸۳] استفاده شده است. همچنین نگاشت حروف فارسی به لاتین مطابق [Megerdooomian 2003] می‌باشد.

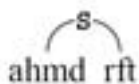
۱-۴ {S(فاعل، فعل)}

در ساده‌ترین الگو زبان فارسی، دو نقش فعل و فاعل وجود دارد که در آن فاعل قبل از فعل قرار می‌گیرد. برای اینکار می‌توان دو ماکرو برای اسامی و افعال تعریف کرد. این دو از طریق پیوند S به یکدیگر متصل می‌شوند.

<noun> : S+;
<verb-it> : S-;

برای مثال، نتیجه پارس جمله (۱) در شکل ۱ آمده است.

(۵۷) احمد رفت.

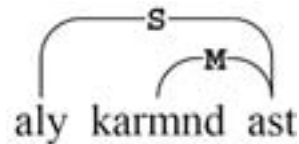


شکل ۱: پارس جمله (۱)

۲-۴ {M(مسند، فعل)}

هرگاه افعال اسنادی در جمله بکار رود، فعل به یک جزء دیگر نیاز دارد که مسند خوانده می‌شود، مسند از طریق پیوند M به فعل متصل می‌شود. خود مسند از طبقه اسامی انتخاب می‌شود. ماکروی <modal> برای تعریف افعال اسنادی بکار رفته است. مثلاً برای جمله (۲) پارس شکل ۲ را خواهیم داشت. همانگونه که مشخص است مسند از طریق پیوند M به فعل مرتبط می‌گردد.

<noun> : S+ or M+;
<modal> : M- & S-;



شکل ۲: پارس جمله (۲)

۳-۴ { (فاعل، مفعول) O, (را، مفعول) OM }

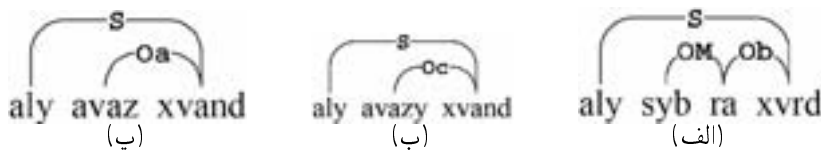
افعال متعدی نیاز به جزء دیگری دارند که مفعول خوانده می‌شوند، مفعول می‌تواند با نشانه را (جمله ۳) و شکل ۳ (الف)، -ی (جمله ۴) و شکل ۴ (ب) یا بدون نشانه (جمله ۵) و شکل ۴ (ج) باشد، با توجه به اینکه -ی واژگ مقید است باید اسمهایی که این نشانه را دارند مشخص نمود، از ماکروی <noun-indef> برای تعریف این اسامی استفاده می‌کنیم. مفعول با پیوند O به فعل متصل می‌شود و نوع آن از طریق زیر نویس تعیین می‌شود.

- <noun> : S+ or M+ or Oa+ or OM+;
- <noun-indef> : Oc+ or OM+;
- ra : OM- & Ob+;
- <verb-t> : {O-} & S-;

(۵۹) علی سیب را خورد.

(۶۰) علی آوازی خواند.

(۶۱) علی آواز خواند.

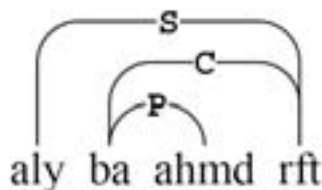


شکل ۳: (الف) مفعول با نشانه را (ب) مفعول با نشانه -ی (ج) مفعول بدون نشانه

۴-۴ { (فعل، متمم) C, (حرف اضافه، متمم) M }

متمم اسمی است که بعد از حرف اضافه آمده و معنی فعل را کامل می‌کند. در جملاتی که فعل آنها لازم است متمم قبل از فعل قرار می‌گیرد (جمله ۶) و شکل ۴). نشانه متمم (حروف اضافه) از طریق پیوند P به اسم متصل شده و از طریق پیوند C به فعل نسبت داده می‌شود.

- <noun> : S+ or M+ or Oa+ or OM+ or P-;
- <prep> : P+ & C+;



شکل ۴: متمم برای افعال لازم

<verb-it> : {C-} & S-;

متمم در جملاتی که مفعول دارند نسبت به آن جایگاه‌های متفاوتی می‌تواند اتخاذ کند [انوری و گیوی ۱۳۸۳].

- اگر مفعول نشانه را داشته باشد، قبل از متمم می‌آید (جمله (۷) و شکل ۵(الف)).
 - اگر مفعول با -ی همراه باشد، پیش و پس از آن (جمله (۸) شکل ۵(ب) و جمله (۹) شکل ۵(پ)).
 - اگر بدون نشانه باشد پس از متمم (جمله (۱۰) و شکل ۵(ت)).
- ماکروی <verb-t> برای اصلاح تعریف افعال متعدی، بازنویسی می‌شود.
- <verb-t> : ({T-} & {C-} & Ob-) or (Oa- & C-) or (Oc- & {C-}) or ({C-} & Oc-) & S-;;

۵-۴ {T (فعل، تمییز)}

تمییز با برخی افعال بکار می‌رود (مثل دانستن) و معنی آن را کامل می‌کند که از آن به عنوان نوعی متمم یا متمم مسندی مفعول [مشکوٰۃ‌الدینی ۱۳۷۹] یاد می‌شود. جایگاه آن قبل از فعل و بعد از مفعول با نشانه را است و از طریق پیوند T به فعل متصل می‌شود. از ماکروی <verb-ic> برای تعریف این افعال استفاده می‌کنیم، مثلاً در جمله (۱۱) مرد نقش تمییزی دارد (شکل ۶).

<noun> : S+ or M+ or Oa+ or OM+ or P- or T+;

<verb-ic> : (({C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & S-;;

۶-۴ وابسته‌های اسم: {B (مضاف‌الیه، مضاف), J (صفت، موصوف)}

این وابسته‌ها شامل صفت و مضاف‌الیه هستند. صفت از طریق پیوند J و مضاف‌الیه از طریق پیوند B به اسم متصل می‌شود (جمله (۱۲) و شکل ۷). برای سادگی پیوندهای اسم را در دو دسته قرار می‌دهیم. پیوندهایی که در آن اسم وابسته به فعل باشد (فاعل، مفعول، متمم یا تمییز) که به کمک ماکروی <noun-vc> خلاصه می‌شوند و غیر آن (مضاف‌الیه و متمم) که ماکروی <noun-nc> به همین منظور بکار می‌رود. اعمال مشابهی هم برای صفات صورت می‌گیرد.

دومین کارگاه پژوهشی زبان فارسی و رایانه - ششم و هفتم تیرماه ۱۳۸۵ / ۱۱۵

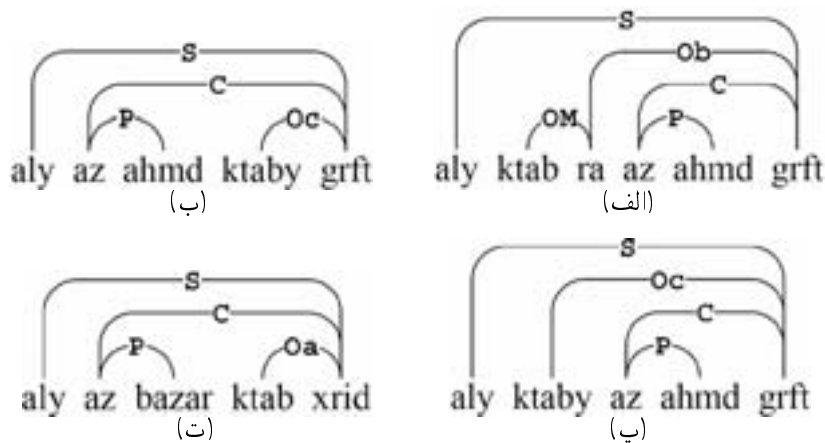
* بازنمایی دانش دستوری زبان فارسی به کمک گرامر پیوندی

(۶۳) علی کتاب را از احمد گرفت.

(۶۴) علی از احمد کتابی گرفت.

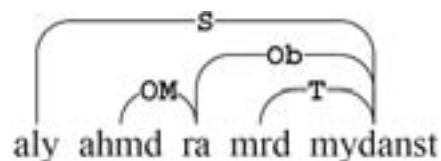
(۶۵) علی کتابی از احمد گرفت.

(۶۶) علی از بازار کتاب خرید.



شکل ۵: جایگاه متعددی در جملات با فعل متعددی (الف) مفعول با نشانهٔ را (ب) و (پ) مفعول با نشانهٔ ی (ت) مفعول بدون نشانه

(۶۷) علی احمد را مرد نمیدانست.



شکل ۶: نقش تمییزی برای افعال ناقص (جمله (۱۱))

<noun-nc> : (B- or P-) & ({@J+} & {B+});

<noun-vc> : {@J+} & {B+} & (S+ or M+ or Oa+ or OM+ or T+);

<noun> : <noun-nc> or <noun-vc>;

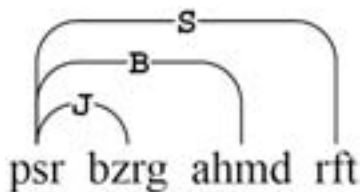
<adj-jc> : P- & ({@J+} & {B+});

<adj-vc> : {@J+} & {B+} & (S+ or M+ or Oa+ or OM+ or T+);

<adj> : <adj-jc> or <adj-vc> or J-;

تعریف صفات شباهت زیادی به اسامی دارد و دلیل آن اینست که صفات می‌توانند جایگاه اسم را اتخاذ کنند، تفاوت آنها در این است که صفات در جایگاه صفت (تشکیل پیوند) J نمی‌توانند همزمان صفت یا مضاف الیه داشته باشند.

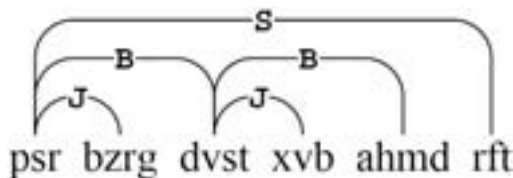
(۶۸) پسر بزرگ احمد رفت.



شکل ۷: وابسته‌های اسم (جمله ۱۱۲)

باید توجه داشت که گرامر پیوندی بخوبی خاصیت بازگشتی بودن زبان را پوشش می‌دهند، مثلاً مضاف الیه خود می‌تواند هسته یک گروه اسمی دیگر قرار گیرد، برای مثال در جمله (۱۳) دوست علاوه بر اینکه نقش مضاف الیه‌ی دارد خود صفت (خوب) و مضاف الیه (احمد) را اختیار کرده است (شکل ۸)، البته تکرار بیش از سه سطح برای زبان فارسی متداول نیست [باطنی ۱۳۸۳].

(۶۹) پسر بزرگ دوست خوب احمد رفت.



شکل ۸: بازگشتی بودن وابسته‌های اسم و مدلسازی آنها به کمک گرامر پیوندی

البته در ساختارهای پیچیده‌تر، ابهام زیادی رخ می‌دهد، مثلاً در جمله (۱۴) دو پارس برای جمله وجود دارد که در مشخص است.

(۷۰) علی زرنگ کلاس را دید.



شکل ۹: ابهام در هنگام متوالی گشتن اسامی در نقش مضافی و صفت (جمله ۱۱۴)

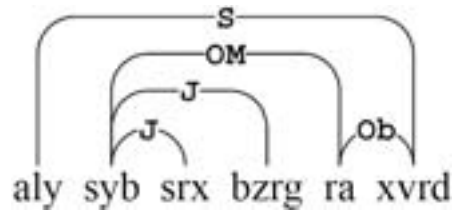
وجود @J- برای اینست که یک اسم ممکن چندین صفت را همزمان اتخاذ کند (جمله ۱۵) و شکل ۱۰).

۷-۴ { (فعل، قید) A }

قید جایگاه ثابتی در جمله ندارد اما با توجه به نوع آن می‌توان جایگاهی را که معمولاً اتخاذ می‌کند پیش بینی کرد. از این حیث، قیود را می‌توان به سه دسته تقسیم کرد [انوری و گیوی ۱۳۸۳]

دومین کارگاه پژوهشی زبان فارسی و رایانه - ششم و هفتم تیرماه ۱۳۸۵ / ۱۱۷

(۷۱) علی سیب سرخ بزرگ را خورد.



شکل ۱۵: صفات متوالی برای یک اسم (جمله ۱۵)

۱-۷-۴ قیدهایی که پیش از فاعل قرار می‌گیرند (جمله ۱۶) و شکل ۱۱ الف).

این دسته شامل قیدهایی تمنا (کاش)، تاسف (افسوس) و تحسین و مسرت (خوشا) می‌باشد که از طریق پیوند Aa به فعل متصل می‌شوند.

(۷۲) کاش هوا گرم بود

<adv-a> : Aa+;

۲-۷-۴ قیدهایی که پس از فاعل قرار می‌گیرند (جمله ۱۷) و شکل ۱۱ ب).

این دسته شامل قیدهایی مکان (هرجا)، کمیت (کمی)، کیفیت (زیبا)، حالت (پیاده)، پرسش (کجا)، تصدیق (حتما) و ترتیب (دوبدو) می‌باشد که از طریق پیوند Ab به فعل متصل می‌شوند.

(۷۳) علی پیاده آمد.

<adv-b> : Ab+;

۳-۷-۴ قیدهایی که هم پیش و هم پس از فاعل قرار می‌گیرند (جملات ۱۸) و (۱۹) و شکل ۱۱ پ و ت).

این دسته شامل قیدهایی زمان (دیروز)، نفی (هرگز) و تردید (شاید) می‌باشد که از طریق پیوند A به فعل متصل می‌شوند، باید توجه داشت که اتصال‌لگر A قابلیت اتصال با هر دو نوع Aa و Ab را دارد (رک. قانون تطابق بخش ۳).

<adv-c> : A+;

به این ترتیب می‌توان ماکروی تعریف افعال را دوباره نوشت، توجه کنید که وجود @ برای پوشش وجود قیدهایی همزمان است.

<modal> : M- & (@Ab- & S- & @Aa-);

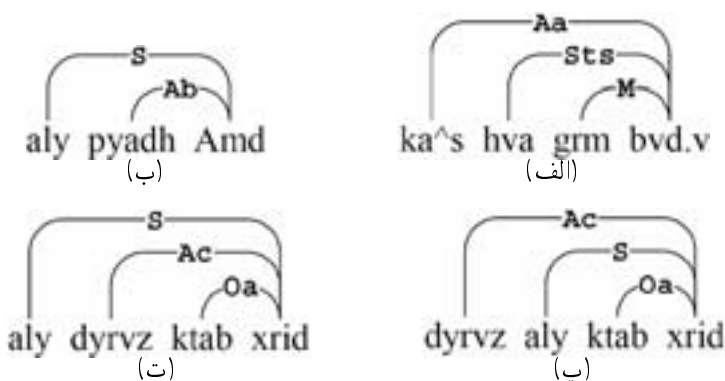
<verb-it> : {C-} & (@Ab- & S- & @Aa-);

<verb-ic> : (({T-} & {C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & (@Ab- & S- & @Aa-);;

<verb-t> : (({C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & (@Ab- & S- & @Aa-);;

(۷۴) دیروز علی کتاب خرید.

(۷۵) علی دیروز کتاب خرید.



شکل ۱۱: جایگاههای مختلف قید (الف) پیش از فاعل (ب) پس از فاعل (ب) و (ت) پیش و پس از فاعل

۸-۴ گوناگونیهای گروه فعلی

با اصلاح تعاریفی که برای فعل بکار رفت، می‌توان انواع گروههای فعلی را که در زبان فارسی وجود دارد، تحت پوشش قرار داد. اگر ساختمان فعل را در حالت کلی بصورت زیر در نظر بگیریم

رابطه ۱ فعل ناقص + عنصر غیر فعلی + فعل واژگانی + عنصر سازنده مجهول + عنصر سازنده حالت با در نظر گرفتن فعل واژگانی به عنوان هسته این ساختار، می‌توان همه دستگاہهای همزمان گروه فعلی را ساخت. فعل واژگانی می‌تواند از یکی از ۱۰ صورت مختلف فعل انتخاب شود، برای سادگی ما تنها صورت ماضی مطلق (در وجه معلوم و مجهول) را در نظر می‌گیریم، تعریف سایر صورتهای به شکل مشابهی انجام می‌گیرد. با در نظر گرفتن تنوعهای مختلف در ساختار فعل، از فرم ماضی مطلق یک فعل مثل خورد، فعلهای می‌توان رفت، می‌شد رفت، بشود رفت، خواهد رفت و رفت ساخته می‌شود. هریک از اعضای رابطه ۱ را می‌توان بصورت زیر تعریف کرد.

۱-۸-۴ {GXY (فعل، فعل ناقص)}

این طبقه به سه زیر طبقه تقسیم می‌شود که هریک را مستقلاً تعریف می‌کنیم. پیوند عمومی مورد استفاده برای آنها GXY است که در آن X زیر طبقه را مشخص می‌کند و Y با توجه به کلمه، ویژگیهای اضافه (عموماً زمان) را کد می‌کند. هریک را بصورت جداگانه تعریف می‌کنیم.

الف) آنهایی که فاقد شخص و عدد هستند.

bayd mytvan my^svd : GAP+;
 my^sd : GAT+;
 btvan b^svd : GAS+;

در تعاریف فوق، P، T و S به ترتیب نشان دهنده زمانهای حال، گذشته و التزامی هستند.

ب) فعل داشتن که قابل صرف برای صیغه‌های مختلف می‌باشد.

dast : GBP+;

ج) فعل خواستن که فقط قابل صرف برای صیغه‌های حال می‌باشد.

xahd : GCP+;

۲-۸-۴ {فعل، عنصر غیر فعلی} K

این طبقه شامل واژه‌هایی است که قبل از فعل قرار می‌گیرند و جزء طبقه افعال نیستند. همه آنها از طریق پیوند K به فعل متصل می‌شوند.

baz va ps ... : K+;

۳-۸-۴ {عنصر ساختمانی حالت، فعل} RT

این طبقه شامل همه صیغه‌های فعل بودن است که علاوه بر نشان دهنده حالت بودن، می‌تواند مستقلاً نیز بکار رود. این کلمه از طریق پیوند R به فعل متصل می‌شود.

bvd : RT+;

۴-۸-۴ دستگاه‌های مختلف فعلی

همانطور که ذکر شد، در این بخش صیغه‌های مختلفی را که از فرم ماضی مطلق ساخته می‌شود، در نظر می‌گیریم. ساخت سایر صیغه‌ها به شکل مشابهی صورت می‌گیرد. برای هر یک از این صیغه‌ها، ماکروبی تعریف می‌شود که در بر دارنده تعریف آن صیغه می‌باشد.

(آ) التزامی غیر شخص (جمله (۲۰))

<ISUB> : {K-} & GAS-;

ب) گذشته اخباری غیر شخصی (جملات (۲۱) و (۲۲))

<IPIN> : {K-} & (GAP- or GAT-);

در فرمهای غیر شخصی جایگاه فاعل خالی می‌ماند در نتیجه تعاریف قبلی (<modal> و غیره) باید اصلاح شود و اتصالگر S حذف شود همچنین با توجه به عدم وجود فاعل اتصالگر قید به A تغییر می‌یابد چرا که در این حالت بین انواع قیدی که در بخش ذکر شد تفاوتی وجود ندارد.

<modal-im> : M- & (@A-);

<verb-it-im> : {C-} & (@A-);

- <verb-ic-im> : (({T-} & {C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & ({@A-});;
<verb-t-im> : (({C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & ({@A-});;

پ) اخباری آینده (جمله ۲۳))

- <FUT> : GCP- & {K-};

ت) گذشته غیر استمراری مطلق (جمله ۲۴))

- <NCSP> : {K-};

و در نهایت ماکروبی برای تعریف افعال ماضی مطلق. ماکروی زیر مانند قبل به ترتیب برای تعریف فرم ماضی مطلق افعال اسنادی، لازم، نیازمند تمییز و متعدی بکار می‌رود.

- <SP-modal> : ((<ISUB> or <IPIN>) & <modal-im>) or ((<FUT> or <NCSP>) & <modal>);
<SP-it> : ((<ISUB> or <IPIN>) & <verb-it-im>) or ((<FUT> or <NCSP>) & <verb-it>);;
<SP-ic> : ((<ISUB> or <IPIN>) & <verb-ic-im>) or ((<FUT> or <NCSP>) & <verb-ic>);;
<SP-t> : ((<ISUB> or <IPIN>) & <verb-t-im>) or ((<FUT> or <NCSP>) & <verb-t>);;

استراتژی پیوند بندی بکار رفته در تعاریف فوق را می‌توان از طریق مثالهای زیر نمایش داد (شکل ۱۲).

۵-۸-۴ وجه مجهولی: { (سازنده مجهول، صفت مفعولی) L }

در وجه مجهولی، همه تغییرات لازم در زمانهای مختلف باید به عنصر سازنده مجهول منتقل شود. برای مثال همان صورت ماضی مطلق فعل شدن را در نظر می‌گیریم. عنصر سازنده مجهول (فعل شدن) همانند سایر افعال تعریف می‌شود با این تفاوت که یک پیوند L با اسم مفعول تشکیل می‌دهد و فاقد مفعول می‌باشند. ملزومات پیوندی افعال متعدی و نیازمند تمییز را در در حالت‌های شخصی و غیر شخصی در ماکروهای <verb-t>، <verb-ic>، <verb-t-im> و <verb-ic-im> تعریف کرده بودیم، ماکروهای زیر همین وظیفه را در حالت مجهولی بر عهده دارند. برای تمایز بین انواع صفت مفعولیهایی که به فعل شد متصل شده‌اند (متعدی یا نیازمند تمییز)، به ترتیب از دو نوع پیوند متفاوت Lt و Lic برای اتصال به فعل شد استفاده می‌کنیم. دو ماکروی <PP-t> و <PP-ic> هم به ترتیب برای تعریف این دو نوع صفت مفعولی استفاده می‌کنیم. دو ماکروی <SP-t-p> و <SP-ic-p> نیز مانند حالت معلوم برای تعریف انواع فرمها و زمانهای فعل مجهولی که از اتصال دو نوع صفت مفعولی مذکور به فعل شد ساخته می‌شود بکار می‌رود. تنها تفاوت اینجا اینست که این ماکروها قبلا در تعریف فرم ماضی مطلق فعل بکار می‌رفتند و در اینجا در تعریف فرم ماضی مطلق عنصر سازنده مجهول. همچنین باز باید توجه کرد که ماکروی <FUT> در اینجا قابل استفاده نیست چرا که

* بازنمایی دانش دستوری زبان فارسی به کمک گرامر پیوندی

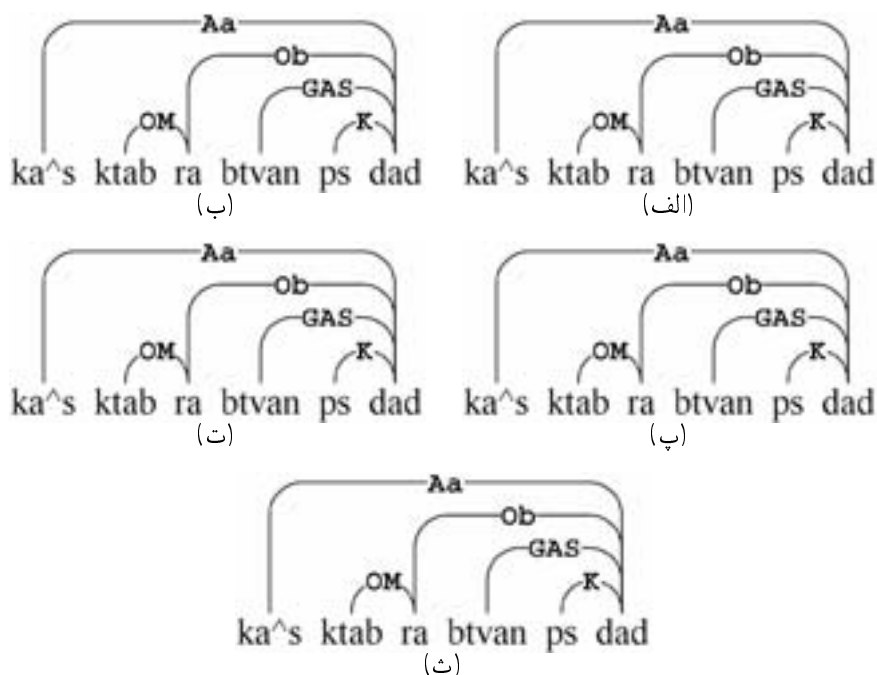
(۷۶) کاش کتاب را بتوان پس داد.

(۷۷) همیشه کتاب را می‌توان پس داد.

(۷۸) دیروز کتاب را می‌شد پس داد.

(۷۹) احمد فردا کتاب را به کتابخانه پس خواهد داد.

(۸۰) احمد کتاب را پس داد.



شکل ۱۲: پیوند بندی حاصل برای جملات (۲۰) تا (۲۴) (به ترتیب الف) تا (ث)

فعل ناقص در حالت مجهول بین فعل واژگانی و عنصر سازنده مجهول قرار می‌گیرد (کشته خواهد شد صحیح است نه خواهد کشته شد) و این ماکرو بصورت دو ماکروی <FUT-t-p> و <FUT-ic-p> بازنویسی شده است که t و ic مانند قبل به نوع فعل دلالت می‌کنند.

- <PP-t> : Lt+;
- <PP-ic> : Lic+;
- <verb-t-p> : {C-} & {@Ab-} & S- & {@Aa-};
- <verb-ic-p> : {T-} & {C-} & {@Ab-} & S- & {@Aa-};
- <verb-ic-im-p> : {T-} & {C-} & {@A-};
- <verb-t-im-p> : {C-} & {@A-};
- <FUT-t-p> : GCP- & Lt- & {K-};
- <FUT-ic-p> : GCP- & Lic- & {K-};

- <SP-t-p> : (Lt- & (((<ISUB> or <IPIN>) & <verb-t-im-p>) or (<NCSP> & <verb-t -p>))) or (<FUT-t-p> & <verb-t-p>);;
- <SP-ic-p> : (Lic- & (((<ISUB> or <IPIN>) & <verb-ic-im-p>) or (<NCSP> & <verb-ic -p>))) or (<FUT-ic-p> & <verb-ic -p>);;
- ^sd : <SP-t -p> or <SP-ic-p>;

همچنین باید توجه داشت که شد خود می‌تواند به‌عنوان یک فعل واژگانی مستقل بکار رود، به همین دلیل می‌توان به تعریف آن ماکروی <SP-modal> را اضافه کرد. این کار می‌تواند در مدخل دیگری از دیکشنری انجام شود.

مثالهای (۲۵) تا (۲۹) می‌تواند در توجیه تعاریف فوق مفید باشد. بالاخص مثال (۲۵) نمونه ای از یک جمله نسبتاً پیچیده برای یک فعل نیازمند تمییز است (شکل ۱۳).

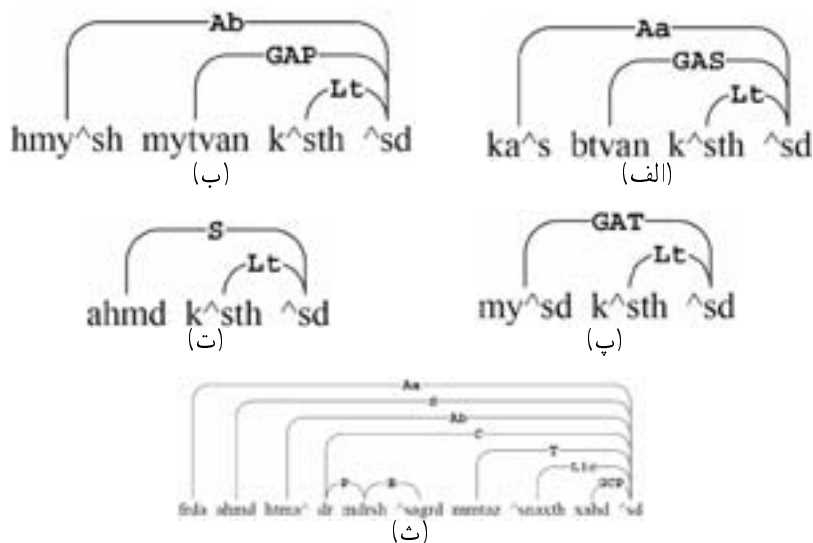
(۸۱) کاش بتوان کشته شد.

(۸۲) همیشه می‌توان کشته شد.

(۸۳) میشد کشته شد.

(۸۴) احمد کشته شد.

(۸۵) فردا احمد حتماً در مدرسه شاگرد ممتاز شناخته خواهد شد..



شکل ۱۳: پیوند بندی حاصل برای جملات (۲۵) تا (۲۹) (به ترتیب الف) تا (ث)

۹-۴ مطابقه فعل و فاعل

در این بخش و بخش بعد، تعریف اسامی و افعال را اندکی پیچیده‌تر می‌کنیم تا نحوه ساخت انواع فعل در زبان فارسی و همچنین نحوه مدیریت تطابق بین فاعل و فعل در شخص و تعداد را تحت پوشش قرار دهیم. با توجه به وابستگی قوانین تطابق در زبان فارسی به واژگان، خاصیت لغوی بودن گرامر پیوندی در اینجا کمک شایانی می‌کند و بسادگی قادر به پشتیبانی آن می‌باشد. برای پوشش دادن دستگاه مطابقه زبان فارسی، نیاز است که ماکرویی که برای تعریف اسامی در نظر گرفته بودیم یعنی $\langle \text{noun} \rangle$ (در اصل $\langle \text{noun-vc} \rangle$)، بسط داده شود، چرا که تطابق در زبان فارسی برای اسامی بی جان، جاندار و انسان متفاوت می‌باشد و همچنین تعاریف را باید برای حالات متفاوت اشخاص و مفرد و جمع بودن هریک، بازنویسی کرد. در ادامه هریک از موارد فوق را بررسی می‌کنیم.

۱-۹-۴ فاعل بی جان

در این حالت فاعل مفرد باید با فعل مطابق باشد اما در حالت جمع در تعداد فاعل اجبار وجود ندارد (جملات (۳۰) و (۳۱)). دو ماکروی $\langle \text{noun-in-s} \rangle$ و $\langle \text{noun-in-p} \rangle$ برای تعریف این اسامی به ترتیب در حالت‌های مفرد و جمع بکار می‌رود.

(۸۶) برگها می‌ریزند.

(۸۷) برگها می‌ریزد.

$\langle \text{noun-vc-in-s} \rangle : \{ @J+ \} \& \{ B+ \} \& (\text{Sts+ or M+ or Oa+ or OM+ or T+});$

$\langle \text{noun-vc-in-p} \rangle : \{ @J+ \} \& \{ B+ \} \& (\text{St+ or M+ or Oa+ or OM+ or T+});$

$\langle \text{noun-in-s} \rangle : \langle \text{noun-nc} \rangle \text{ or } \langle \text{noun-vc-in-s} \rangle ;$

$\langle \text{noun-in-p} \rangle : \langle \text{noun-nc} \rangle \text{ or } \langle \text{noun-vc-in-p} \rangle ;$

همانطور که مشخص است تنها پیوندی که تغییر کرده پیوند مابین فاعل و فعل است (S)، زیر نویسه‌های t و s برای نشان دادن سوم شخص و مفرد بودن بکار می‌روند. همانگونه که مشاهده می‌شود، مطابقت در گرامر پیوندی عموماً از طریق زیرنویسها انجام می‌گیرد، دلیل این کار با توجه به نحوه تطابق پیوند ها که در بخش ذکر شد، منطقی بنظر می‌رسد.

۲-۹-۴ فاعل جاندار غیر انسان

در این حالت بین فاعل و فعل تطابق وجود دارد (جملات (۳۲) و (۳۳)). $\langle \text{noun-an-s} \rangle$ و $\langle \text{noun-an-p} \rangle$ برای تعریف این اسامی به ترتیب در حالت‌های مفرد و جمع بکار می‌رود.

(۸۸) گربه آمد

(۸۹) گربه‌ها آمدند.

$\langle \text{noun-vc-an-s} \rangle : \{ @J+ \} \& \{ B+ \} \& (\text{Sts+ or M+ or Oa+ or OM+ or T+});$

- <noun-vc-an-p> : { @J+ } & { B+ } & (Stp+ or M+ or Oa+ or OM+ or T+);
 <noun-an-s> : <noun-nc> or <noun-vc-an-s>;
 <noun-an-p> : <noun-nc> or <noun-vc-an-p>;

۳-۹-۴ مسند الیہ انسان

در همه حالات مطابقت وجود دارد جز در حالتی که مسندالیہ سوم شخص مفرد باشد، در این حالت در تعداد مطابقت ضروری نیست (جملات (۳۴) و (۳۵)). برای سادگی فقط حالت اول شخص مفرد را تعریف می‌کنیم و تعریف سایر شخصها و تعدادها به صورت مشابه تعریف می‌شود. در نتیجه می‌توان تعریف کرد.

(۹۰) استاد محترم آمد

(۹۱) استاد محترم آمدند.

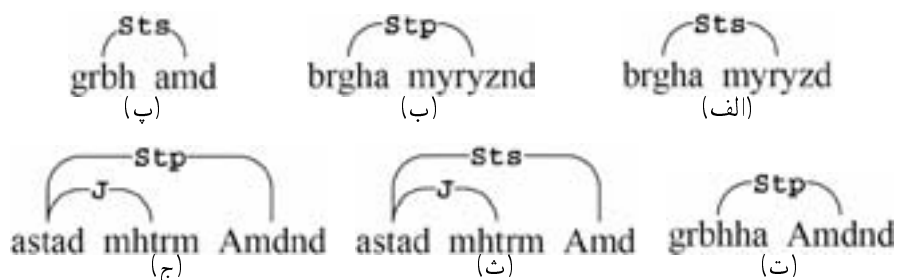
- <noun-vc-hu-fs> : { @J+ } & { B+ } & (Sfs+ or M+ or Oa+ or OM+ or T+);
 <noun-vc-hu-ts> : { @J+ } & { B+ } & (St+ or M+ or Oa+ or OM+ or T+);
 <noun-hu-ts> : <noun-nc> or <noun-vc-hu-fs>;
 <noun-hu-tp> : <noun-nc> or <noun-vc-hu-ts>;

۴-۹-۴ مطابقت افعال

همانطور که گفته شد باید ماکروهای تعریف افعال بازنویسی شود. ما بسط این ماکروها را برای حالت سوم شخص مفرد و فقط برای دو ماکروی <modal> و <verb-it> بررسی می‌کنیم. دیگر ماکروها و همچنین تعریف افعال ناقص (بخش ۴-۸-۱) و عنصر حالت (بخش ۴-۸-۳) به شکل مشابهی اصلاح می‌شوند (رج ضمیمه الف).

- <modal-ts> : M- & ({ @Ab- } & Sts- & { @Aa- });
 <verb-it-ts> : { C- } & ({ @Ab- } & Sts- & { @Aa- });

در نهایت می‌توان صحت پارس قوانین را برای جملات (۳۰) تا (۳۵) سنجید (شکل ۱۴).



شکل ۱۴: تطبیق فاعل و فعل برای جملات (۳۰) تا (۳۵) به ترتیب الف تا ج

۵ نتیجه گیری و کارهای آینده.

از مثالها و مراحل تکمیل گرامر، روشن بنظر می‌رسد که توسعه این گرامر بسادگی صورت می‌گیرد. آنچه در این مقاله بررسی شد، نحوه تعریف واژگان برای تحت پوشش قرار دادن نقشهای اساسی و الگوهای اولیه برای بندهای مهین زبان فارسی می‌باشد. دستور زبان فارسی در سطح جمله دارای الگوهای بسیار پیچیده‌تر و ظریفتر می‌باشد که می‌تواند محور کارهای آینده قرار گیرد، مثلاً می‌توان به جملات مرکب با الگوهای تکرار مختلف (هسته ای یا خوشه ای) اشاره کرد. البته ایجاد بسترهای تست مناسب و استفاده از این گرامرها جهت ادغام در سیستمهای دیگر از اهداف آینده این پروژه می‌باشد.

۶ خلاصه

گرامر پیوندی به عنوان فرمالیسمی برای بازنمایی زبانهای مستقل از متن به علت لغوی بودن و در نتیجه کاهش پیچیدگی دستور زبان و توزیع آن در سطح همه کلمات و نیز مقیاس پذیری آن می‌تواند به عنوان بدیلی برای گرامرهای با ساختار گروهی بکار رود. این گرامر بدون نیاز به مفهوم واحد و تجزیه مستقیم همه خواص یک گرامر برای بازنمایی دانش زبان طبیعی را دارا می‌باشد. هدف ما در این تحقیق بررسی کارایی این گرامرها برای زبان فارسی بود. این کار از طریق تطبیق گرامر پیوندی با افزودن پیوندهای مورد نیاز به آنها انجام شد. ما با آغاز از نقشهای اساسی زبان فارسی (فاعل و فعل و سپس افزودن دیگر نقشها مجموعه ای از ملزومات پیوندی را فراهم آوردیم که قادر به پردازش جملات ساده و ابتدایی زبان باشد. تطبیق فعل و فاعل در زبان فارسی نمونه ای از خواص لغوی بود که گرامر پیوندی با توجه به لغوی بودن به سادگی قادر به مدیریت آنها می‌باشد. در هر مرحله با افزودن مثالهایی سعی نمودیم روش خود را ملموستر نماییم. با تست مرحله به مرحله قوانین به کمک پارسر گرامر پیوندی از مجموعه بدست آمده یک مجموعه سازگار استخراج شد (کل دیکشنری در ضمیمه الف گردآوری شده است).

مراجع

- [انوری و گیوی ۱۳۸۳] انوری، حسن، گیوی، حسن، دستور زبان فارسی ۲، ویرایش ۲، چاپ بیست و سوم، ۱۳۸۳، انتشارات فاطمی
- [آیت ۱۳۸۰] آیت، مریم، یک گرامر محاسباتی برای زبان فارسی، پایان نامه کارشناسی ارشد رشته مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، ۱۳۸۰
- [باطنی ۱۳۸۳] باطنی، محمد رضا، توصیف ساختمان دستوری زبان فارسی، ۱۳۸۳، انتشارات امیرکبیر.
- [خانلری ۱۳۸۲] خانلری، پرویز، دستور زبان فارسی، ویرایش دوم، ۱۳۸۲، انتشارات توس.
- [دبیرمقدم ۱۳۸۳] دبیرمقدم، محمد، زبان شناسی نظری: پیدایش و تکوین دستور زایشی، ویرایش ۲، سازمان مطالعه و تدوین کتب علوم انسانی و دانشگاهها (سمت)، ۱۳۸۳.
- ۱۲۶ / دومین کارگاه پژوهشی زبان فارسی و رایانه - ششم و هفتم تیرماه ۱۳۸۵

[عبدالله زاده ۱۳۸۳] عبدالله زاده، احمد. یادداشتهای درس پردازش زبان طبیعی، دانشگاه صنعتی امیرکبیر، ۱۳۸۳

[فیلی و قاسم ثانی ۱۳۸۳] ه. فیلی و غ. قاسم ثانی، استفاده از گرامر یکسان ساز برای ترجمه انگلیسی به فارسی، در مجموعه مقالات دهمین کنفرانس سالانه انجمن کامپیوتر ایران، جلد اول، ص ۹۴۴-۹۳۶، ۱۳۸۳

[فیلی و قاسم ثانی ۱۳۸۲] ه. فیلی و غ. قاسم ثانی، استفاده از گرامر درخت-افزایشی برای ترجمه انگلیسی به فارسی، در مجموعه مقالات نهمین کنفرانس سالانه انجمن کامپیوتر ایران، جلد اول ص ۶۴۷-۶۳۹، ۱۳۸۲

[مشکوٰۃ الدینی ۱۳۷۹] مشکوٰۃ الدینی. مهدی، دستور زبان فارسی بر اساس نظریه گشتاری، ویرایش دوم، ۱۳۷۹، دانشگاه فردوسی مشهد.

[Aubin 2003] Aubin, S.. Evaluation comparative de deux analyseurs produisant des relations syntaxiques (Comparative evaluation of two analyzers producing syntactic relations). Traitement Automatique des Langues Naturelles 2003

[Allen 1997] Allen, J. (1997). Natural Language Understanding, Allen, J., Chapter VII. Benjamins Cummings ditions, 1997.

[Chomsky 1957] Chomsky, N. (1957). Syntactic Structures. The Hague: Mouton.

[Hashara et al. 2003] Harsha, V. Madhyastha, N. Balakrishnan, & K. R. Ramakrishnan. 2003. Event Information Extraction Using Link Grammar. 13th International WorkShop on Research Issues in Data Engineering: Multi-lingual Information Management (RIDE'03)

[Hunt 1993] Hunt, A.J. Utilizing Prosody to Perform Syntactic Disambiguation. In Proc. Eurospeech 1993, pages 1339-1342. 1993

[Joshi 1985] Joshi. A.K. Tree Adjoining Grammars: How much context sensitivity is required to provide a reasonable structural description. In I. Karttunen D. Dowty and A. Zwicky, editors, Natural Language Parsing, pages 206-250, Cambridge, U.K., Cambridge University Press. 1985.

[Diang et al 2003] Ding, J. Berleant, D. Xu, J and Fulmer, A.W. Extracting Biochemical Interactions from MEDLINE Using a Link Grammar Parser. 15th IEEE International Conference on Tools with Artificial Intelligence (IC-TAI'03). 2003

[Gazdar et al. 1982] Gazdar, Gerald and G. Pullum. Generalized Phrase Structure Grammar: A Theoretical Synopsis. Indiana University Linguistics Club, Bloomington, Indiana. 1982.

[Hudson 1984] Hudson, R., Word Grammar, Basil Blackwell, 1984.

[Lafferty et al. 1992] John Lafferty, Daniel Sleator, and Davy Temperley. Grammatical Trigrams: A Probabilistic Model of Link Grammar. Proceedings of the AAAI conference on Probabilistic Approaches to Natural Language, October, 1992.

[Liu 2002] Carol Liu. Towards A Link Grammar for Chinese. Submitted for publication in Computer Processing of Chinese and Oriental Languages - the Journal of the Chinese Language Computer Society. 2002

- [Megerdoomian 2003] K.Megerdoomian, Unification-based Persian morphology. In proceedings of CICLing 2000, Alexander Gelbukh, Center of Investigation on Computation-IPN, Mexico, 2003
- [Meřćuk 1988] Meřćuk, I. A. Dependency Syntax: Theory and Practice, State University of New York Press 1988.
- [Naur 1958] Naur, P., Revised Report on the Algorithmic Language, ALGOL 60, Communications of the ACM 6(1). January 1958.
- [Pollard 1994] Pollard, C., and Sag I. A., Head-driven Phrase Structure Grammar, Chicago Univ Press, 1994
- [Radford 1981] Radford, A., Transformational Syntax, Cambridge University Press, 1981.
- [Kaplan and Bresnan 1982] Kaplan, R and Bresnan, J. Lexical functional grammar: a formal system for grammatical representation. In J. Bresnan, editor, The Mental Representation of Grammatical Relations, pages 173–281. 1982
- [Russel and Norvig 2003] Russell, S.J. and Norvig, P. Artificial Intelligence: A Modern Approach. Prentice Hall, 2nd edition, °
- [Sleator and Temperley 1991] Sleator, D. and Temperley, D. Parsing English with a Link Grammar. Carnegie Mellon University Computer Science technical report CMU-CS-91-196, October 1991
- [Sleator and Temperley 1993] Daniel Sleator and Davy Temperley. Parsing English with a LinkGrammar. Third International Workshop on Parsing Technologies. 1993.
- [Tanimoto 1987] Tanimoto, S.L. The Elements of Artificial Intelligence. Computer Science Press, Rockville, MD, 1987.

ضمیمه الف. دیکشنری نهایی

در این بخش، خلاصه آنچه گفته شد و همچنین تعریف واژگان مورد استفاده در این مثالها به عنوان دیکشنری نهایی آمده است، این تعاریف توسط پارسر گرامر پیوندی^۱ تست شده است.

<noun-nc>	: (B- or P-) & (@J+) & {B+};;
<noun-vc-in-s>	: @J+ & {B+} & (Sts+ or M+ or Oa+ or OM+ or T+);
<noun-vc-in-p>	: @J+ & {B+} & (St+ or M+ or Oa+ or OM+ or T+);
<noun-in-s>	: <noun-nc> or <noun-vc-in-s>;
<noun-in-p>	: <noun-nc> or <noun-vc-in-p>;
<noun-vc-an-s>	: @J+ & {B+} & (Sts+ or M+ or Oa+ or OM+ or T+);

1) <http://www.link.cs.cmu.edu/link/ftp.html> موجود در آدرس

<noun-vc-an-p>	: { @J+ } & { B+ } & (Stp+ or M+ or Oa+ or OM+ or T+);
<noun-an-s>	: <noun-nc> or <noun-vc-an-s>;
<noun-an-p>	: <noun-nc> or <noun-vc-an-p>;
<noun-vc-hu-fs>	: { @J+ } & { B+ } & (Sfs+ or M+ or Oa+ or OM+ or T+);
<noun-vc-hu-ts>	: { @J+ } & { B+ } & (St+ or M+ or Oa+ or OM+ or T+);
<noun-hu-fs>	: <noun-nc> or <noun-vc-hu-fs>;
<noun-hu-ts>	: <noun-nc> or <noun-vc-hu-ts>;
<noun-indef>	: Oc+ or OM+;
<adj-jc>	: P- & ({ @J+ } & { B+ });
<adj-vc>	: { @J+ } & { B+ } & (S+ or M+ or Oa+ or OM+ or T+);
<adj>	: <adj-jc> or <adj-vc> or J-;
<prep>	: P+ & C+;
ra	: OM- & Ob+;
<adv-a>	: Aa+;
<adv-b>	: Ab+;
<adv-c>	: A+;
<modal-ts>	: M- & ({ @Ab- } & Sts- & { @Aa- });
<modal-tp>	: M- & ({ @Ab- } & Stp- & { @Aa- });
<verb-it-ts>	: { C- } & ({ @Ab- } & Sts- & { @Aa- });
<verb-it-tp>	: { C- } & ({ @Ab- } & Stp- & { @Aa- });
<verb-ic-ts>	: (({ T- } & { C- } & Ob-) or (Oa- & { C- }) or (Oc- & { C- }) or ({ C- } & Oc-)) & ({ @Ab- } & Sts- & { @Aa- });;
<verb-t-ts>	: (({ C- } & Ob-) or (Oa- & { C- }) or (Oc- & { C- }) or ({ C- } & Oc-)) & ({ @Ab- } & Sts- & { @Aa- });;
bayd mytvan my [^] svd.g	: GAP+;
my [^] sd	: GAT+;
btvan bs [^] vd	: GAS+;
da [^] st	: GBPts+;
xahd	: GCPts+;
ps	: K+;
bvd.r	: RTts+;
<ISUB>	: { K- } & GAS-;
<IPIN>	: { K- } & (GAP- or GAT-);

<modal-im>	: M- & {@A-};
<verb-it-im>	: {C-} & {@A-};
<verb-ic-im>	: (({T-} & {C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & {@A-};;
<verb-t-im>	: (({C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & {@A-};;
<FUT-ts>	: GCPts- & {K-};
<NCSP>	: {K-};
<SP-modal-ts>	: ((<ISUB> or <IPIN>) & <modal-im>) or ((<FUT-ts> or <NCSP>) & <modal-ts>);;
<SP-it-ts>	: ((<ISUB> or <IPIN>) & <verb-it-im>) or ((<FUT-ts> or <NCSP>) & <verb-it-ts>);;
<SP-ic-ts>	: ((<ISUB> or <IPIN>) & <verb-ic-im>) or ((<FUT-ts> or <NCSP>) & <verb-ic-ts>);;
<SP-t-ts>	: ((<ISUB> or <IPIN>) & <verb-t-im>) or ((<FUT-ts> or <NCSP>) & <verb-t-ts>);;
<PP-t>	: ;
<PP-ic>	: ;
<verb-t-ts-p>	: {C-} & {@Ab-} & Sts- & {@Aa-};
<verb-ic-ts-p>	: {T-} & {C-} & {@Ab-} & Sts- & {@Aa-};
<verb-ic-im-p>	: {T-} & {C-} & {@A-};
<verb-t-im-p>	: {C-} & {@A-};
<FUT-t-ts-p>	: GCPts- & Lt- & {K-};
<FUT-ic-ts-p>	: GCPts- & Lic- & {K-};
<SP-t-ts-p>	: (Lt- & (((<ISUB> or <IPIN>) & <verb-t-im-p>) or (<NCSP> & <verb-t-ts-p>))) or (<FUT-t-ts-p> & <verb-t-ts-p>);;
<SP-ic-ts-p>	: (Lic- & (((<ISUB> or <IPIN>) & <verb-ic-im-p>) or (<NCSP> & <verb-ic-ts-p>))) or (<FUT-ic-ts-p> & <verb-ic-ts-p>);;
^sd	: ;
hva avaz syb psr ktab bazar	
klas brg ktabxanh mdrsh	: <noun-in-s>;
brgha	: <noun-in-p>;
grbh	: <noun-an-s>;
grbhha	: <noun-an-p>;
^sagrd ahmd karmnd aly mrd	
dvst astad	: <noun-hu-ts>;
ktaby avazy	: <noun-indef>;
dr az ba bh	: <prep>;

mmtaz zrng srx bzrg xvb grm	
mhtrm	: <adj>;
ka ^ˆ s	: <adv-a>;
htma ^ˆ hmy ^ˆ sh pyadh	: <adv-b>;
dyrvz frda	: <adv-c>;
my ^ˆ svd.v bvd.v ast	: <modal-ts>;
my ^ˆ svnd	: ;
mydanst	: <verb-ic-ts>;
myryzd	: <verb-it-ts>;
Amdnd myryznd	: <verb-it-tp>;
dad grft xvrđ xvand xryd	
dyd k ^ˆ st	: <SP-t-ts>;
Amd rft	: <SP-it-ts>;;
^ˆ snaxt	: <SP-ic-ts>;
k ^ˆ sth	: <PP-t>;
^ˆ snaxth	: <PP-ic>;