

تحلیل نحوی زبان فارسی به کمک گرامر پیوندی^۱

آرمین سجادی و احمد عبدالله زاده بارفروش
دانشکده‌ی کامپیوتر، دانشگاه صنعتی امیرکبیر
نویسنده‌ی عهده‌دار مکاتبات: آرمین سجادی

چکیده

بازنمایی دستور زبان را می‌توان به عنوان مهم‌ترین مرحله 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 تجزیه‌ی به واحدهای کوچک‌تر تحلیل نمایند. دیدگاه‌های متفاوتی برای این کار وجود دارد که گرامر پیوندی به عنوان یکی از آن‌ها مینا 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 نحوی عملکرد آن 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، توسعه‌ی کار را مبتنی بر این روابط قرار دهیم و در هر مرحله، تعداد این روابط را افزایش دهیم تا حدی که مجموعه قواعدی برای متون رسمی نوشتاری و ساخت‌های ابتدایی فراهم آوریم. برای اطمینان از انسجام مجموعه، قواعد در هر مرحله به کمک تقطیع‌گر گرامر پیوندی و روی یک بستر مناسب آزمایش شده است.

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

۱- مقدمه

دو دلیل عمده دارد (۱) به اندازه‌ی کافی توانا هستند تا زبان طبیعی را بازنمایی کنند (۲) به اندازه‌ی کافی محدود هستند تا به سادگی تقطیع شوند [۱]. این صورت‌بندی مبتنی بر تجزیه‌ی سلسله‌مراتبی جمله به واحدهای کوچک‌تر می‌باشد و به همین دلیل تحت عنوان تجزیه‌ی سازه‌ی بلافصل^۲ و تقطیع، مشهور است که زبان‌شناسان موسوم به ساخت‌گرای آمریکایی برای توصیف نحو از آن استفاده کردند [۴]. صورت‌های مختلف این نظریه علی‌رغم تنوع‌های گوناگون، همگی در این خصیصه، مشترک هستند و به همین دلیل [2] همه‌ی 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،

آنچه دانش را از مفاهیم مشابهی دیگری چون داده و اطلاعات متمایز می‌سازد، سطح انتزاع در آن‌ها می‌باشد [۱] و دانش دستوری از مهم‌ترین دانش‌هایی است که در هر سیستم پردازش زبان طبیعی وجود دارد. تاکنون صورت‌بندی‌های متعددی برای بازنمایی دستور معرفی شده که مهم‌ترین آن‌ها قواعد بازنویسی^۳ زبان‌های مستقل از متن می‌باشد که به‌طور مستقل در [2] برای بازنمایی زبان طبیعی و در [3] برای استفاده در تقطیع^۴ زبان‌های مصنوعی به‌کار رفت و به BNF^۵ مشهور است. استفاده از این صورت‌بندی،

¹ linkgramer

² Formalism

³ Rewrite Rules

⁴ Parse

⁵ Backus- Naur Form

⁶ Immediate Constituent Analysis

آن‌ها برای کنترل پیچیدگی‌های زبان. اما این نظریه با انتقاداتی نیز همراه شد؛ یکی از آن‌ها [15] و دیگری [10] بود که به جای تکیه بر واحدها، استفاده از نوعی وابستگی ساختاری بین اجزای جمله را پیشنهاد کردند. مثلاً در این روش از یک گراف جهت‌دار مسطح بین کلمات استفاده می‌شد که تشکیل یک درخت را می‌دادند. به‌جز ریشه، همه‌ی کلمات دارای یک یال به سمت فقط و فقط یک کلمه‌ی دیگر داشتند و هیچ یالی به ریشه ختم نمی‌شد. گرامر پیوندی (LG) علاوه بر لغوی بودن هر سه شرط مطرح شده برای گرامرها را دارد [16]. به هر حال گرامر پیوندی را باید، ادامه‌ی این تلاش‌ها و راهی برای استفاده از خود کلمات به جای واحدها 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 صورت‌بندی جدید برای بازنمایی زبان‌های مستقل از متن می‌باشد [17]. اما برای مقایسه‌ی بین دو صورت‌بندی مختلف نیز ملاک‌هایی موجود است، از جمله [16].

- سادگی در بازنمایی
- کارایی عملیات تقطیع
- مفید بودن خروجی تقطیع‌گر برای پردازش‌های بعدی
- به این ترتیب می‌توان مشخصه‌های گرامر پیوندی را به‌صورت زیر بر شمرد:
- ساده‌سازی گرامر
- امکان ارائه‌ی یک مدل احتمالاتی [18].
- برخلاف گرامرهای مبتنی بر ساختار گروهی، کلماتی که از لحاظ نحوی یا معنایی به‌هم مرتبط هستند، بعد از تقطیع به هم پیوند می‌شوند و این مسأله 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 پیوندی این است که علی‌رغم قائل نبودن به واحدها، اطلاعات مربوط به آن‌ها به‌سادگی قابل استخراج می‌باشد.
- گرامرهای پیوندی در سطح وسیعی برای زبان انگلیسی به‌کار رفته‌اند. تعریف کاملی از واژه‌نامه برای بازنمایی بخش وسیعی از زبان انگلیسی در [17] آمده است. از گرامر پیوندی برای استخراج رویدادها [19]، استخراج دانش از سیستم‌های اطلاعاتی پزشکی [20] و موارد زیاد دیگری استفاده شده است. همچنین در سطح زبان‌های دیگر نیز مورد توجه قرار

گرامرهای با ساخت گروهی تعمیم یافته [5] (GPSG) ¹ و نوع متأخر و بسیار پیچیده‌تری از آن با عنوان گرامرهای با ساخت گروهی هسته-بنیان ² (HPSG) [6] می‌باشد. از سال ۱۹۸۰ به بعد، بیشتر تلاش‌ها به سمت بازنمایی‌های با واژگان پیچیده‌تر و گرامرهای کمتر هدایت شده است [7]. اولین دسته از این گرامرهای مبتنی بر لغت، گرامر نقشی لغوی ³ [8] می‌باشد. از دیگر گرامرهای مبتنی بر این مفهوم می‌توان به گرامر درخت/افزایشی ⁴ (TAG) [9] 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 [10] و گرامر پیوندی است. تاکنون در سیستم‌های پردازش زبان فارسی روش‌هایی چون قواعد بازنویسی مستقل از متن [۱۱]، گرامرهای درخت-افزایشی [۱۲] و HPSGها [۱۳] به‌کار گرفته شده ولی استفاده از گرامرهای لغوی سابقه‌ی 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. در بخش دو این گرامرها را به‌صورت مختصری معرفی می‌کنیم و با ترمینولوژی آن‌ها در بخش سه آشنا می‌شویم. در بخش چهار با شروع از نقش‌های اولیه و الگوهای ساده در هر مرحله سعی در پیچیده‌تر کردن این تعاریف می‌کنیم تا در نهایت یک واژه‌نامه‌ی مختصر برای پردازش متون رسمی نوشتاری و ساخت‌های ابتدایی ⁵ زبان فارسی ارائه دهیم. طبیعتاً سیستم پیاده‌سازی شده واقعی، دارای مجموعه قواعد بسیار بیشتری است اما هدف اصلی مقاله تشریح رهیافت کلی می‌باشد. یک نکته قابل توجه، مقیاس‌پذیری ⁶ مجموعه قواعد است و به همین دلیل از هر دسته قواعد تنها نمونه‌های کوچکی مطرح شده است.

۲- معرفی گرامر پیوندی

[14] مهم‌ترین مزیت گرامرهای مبتنی بر واحد را سه مورد عنوان می‌کند: اول تحت پوشش قرار دادن نظم توزیعی ⁷ زبان، دوم قابلیت بازگشتی بودن آن‌ها و سوم مقید بودن

¹ Generalized Phrase Structure Grammar

² Head driven Phrase Structure Grammar

³ Lexical Functional Grammars

⁴ Tree Adjoining Grammar

⁵ Trivial

⁶ Scalability

⁷ Distributional regularities

می‌شوند که اسامی آن‌ها با هم مطابق بوده و اتصال گر + (-) در سمت چپ (راست) پیوند قرار گیرد.

• **ملزومات پیوندی**^{۱۱}. تعریف پیوندها برای یک اتصال گر به شکل فرمول‌های خوش فرمی است که توسط عمل‌گرهای گرامر پیوندی ساخته می‌شود که فهرست آن‌ها در جدول (۱) آمده است.

• **تطابق پیوندی**^{۱۲}. در حالت کلی یک اتصال گر با یک یا بیشتر حرف بزرگ شروع می‌شود که نام آن را مشخص می‌کند و با دنباله‌ای از حروف کوچک ادامه می‌یابد که زیرنویس نامیده می‌شود. قاعده‌ی کلی تطابق بین دو اتصال گر را می‌توان به صورت زیر خلاصه کرد:
 دو اتصال گر $Ax_1...x_m$ و $By_1...y_n$ قابلیت اتصال دارند هرگاه (الف) A با B مطابقت کند و (ب) $x_1...x_m$ با $y_1...y_n$ مطابقت کند و $m \geq n$ یا $x_1...x_m$ با $y_1...y_n$ مطابقت کند و $m < n$.

• **تعریف**^{۱۳}. به مجموعه قواعد یک کلمه، تعریف آن گفته می‌شوند.

• **واژه‌نامه**^{۱۴}. به مجموعه‌ای که حاوی کلمات و تعاریف آن‌ها می‌باشد، واژه‌نامه گفته می‌شود که با تلقی متداول از آن تفاوت دارد.

• **زیرنویس کلمات**^{۱۵}. کلماتی که دارای دو تعریف مختلف باشند، باید با زیرنویس‌های مخالف جدا شوند. زیرنویس حروف کوچکی است که بعد از نقطه قرار می‌گیرد، مانند *weekly.e*.

جدول ۱. عمل‌گرهای گرامر پیوندی

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

گرفته است، برای نمونه می‌توان به زبان فرانسوی [21] و چینی [22] اشاره کرد. هم‌چنین گرامر پیوندی احتمالاتی^۱ نیز توانایی زیادی در زمینه‌ی پردازش گفتار نشان داده است [23] که ارزیابی عملکرد آن‌ها برای زبان فارسی، هدف این مقاله می‌باشد.

۳- نمادگذاری گرامرهای پیوندی

گرامر پیوندی به عنوان یکی از گرامرهای وابستگی توسط *دانیل اسلیتور*^۲ معرفی شد [16] و [17]. هر گرامر پیوندی شامل مجموعه‌ای از کلمات (نمادهای پایانی^۳) است که هر یک دارای تعدادی ملزوم پیوندی هستند. هر جمله که توسط این گرامر مورد قبول قرار می‌گیرد، اگر بتوان بین کلمات یال‌هایی کشید به طوری که این یال‌ها ملزومات پیوندی کلمات را ارضا کنند و قواعد زیر (که فراقواعد^۴ نامیده می‌شوند) را رعایت کنند.

• **مسطح بودن**^۵. پیوندها هنگامی که از بالای کلمات وصل می‌شوند، نباید هم‌دیگر را قطع کنند.

• **هم بندی**^۶. پیوندها همه‌ی کلمات را به هم وصل کنند (گراف حاصل هم بند باشد).

• **ترتیب**^۷: هنگامی که اتصال‌گرهای یک کلمه از چپ به راست پیمایش می‌شوند، کلماتی که به آن‌ها متصل شده‌اند از نزدیک به دور تغییر می‌یابند.

• **انحصار**^۸. هیچ دو پیوندی یک جفت کلمه را به هم متصل نمی‌کند.

دیگر مفاهیم گرامر پیوندی که به آن‌ها رجوع می‌شود به طور خلاصه عبارتند از:

• **پیوند**^۹. پیوندها جفت کلمات را به هم مرتبط می‌کنند که به هر یک اسمی اختصاص داده می‌شود.

• **اتصال گر**^{۱۰}. پیوندها از طریق اتصال‌گرهای کلمات به آن‌ها متصل می‌شوند که از آن به عنوان ارضا کردن اتصال گر ذکر می‌شود. هر اتصال گر شامل یک نام پیوند به همراه + یا - می‌باشد، دو اتصال گر هنگامی که هم متصل

¹ Probabilistic Link-Grammars

² Daniel Sleator

³ Terminal

⁴ Meta Rules

⁵ Planarity

⁶ Connectivity

⁷ Ordering

⁸ Exclusion

⁹ Link

¹⁰ Connector

¹¹ Linking Requirements

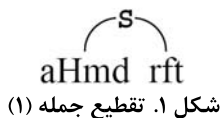
¹² Link Matching

¹³ Definition

¹⁴ Dictionary

¹⁵ Subscripts

برای مثال، نتیجه‌ی تقطیع جمله‌ی (۱) در شکل ۱ آمده است.
(۱) احمد رفت.

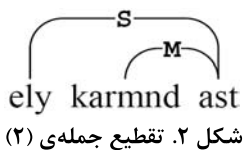


۴-۲- {M (فعل، مسند)}

هرگاه افعال اسنادی در جمله به کار رود، فعل به یک جزء دیگر نیاز دارد که مسند خوانده می‌شود، مسند از طریق پیوند M به فعل متصل می‌شود. خود مسند از طبقه‌ی اسامی (یا صفات، رج. بخش ۴-۶) انتخاب می‌شود. ماکروی <modal> برای تعریف افعال اسنادی به کار رفته است. مثلاً برای جمله‌ی (۲) تقطیع 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-;

(۲) علی کارمند است.



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

افعال متعدی، نیاز به جزء دیگری دارند که مفعول خوانده می‌شوند؛ مفعول می‌تواند با نشانه‌ی [جمله‌ی ۰ و شکل ۳ الف]، "ی" [جمله‌ی (۴) و شکل ۳ ب]، بدون نشانه‌ی [جمله‌ی (۵) و شکل ۳ پ] و یا هردو [جمله‌ی (۶) و شکل ۳ ت] 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 زیرنویس تعیین می‌شود.

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

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 این کار مورد سنجش قرار دهیم؛ تلاش ما بر این بوده است که توسعه‌ی کار، قدم به قدم صورت گیرد تا معیاری جهت ارزیابی مقیاس‌پذیری گرامر پیوندی به دست دهد. مبتنی بودن بیشتر مطالعات سنتی دستور زبان فارسی بر رابطه‌ها از یک سو و نقش محوری و مرکزی رابطه‌ها در دستور زبان که اساس کار دستوره‌های رابطه‌ای را تشکیل می‌دهد [۴] از سوی دیگر، گرامر پیوندی را به‌عنوان صورت‌بندی‌ای که در آن رابطه‌ها به‌سادگی از نتیجه‌ی تقطیع، قابل استخراج هستند، مناسب می‌نمایند. ما نیز در این مقاله بر این رابطه‌ها تکیه کرده‌ایم و با شروع از مجموعه‌ی اولیه‌ی {فاعل، فعل} در هر مرحله یک یا چند رابطه 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 $X(\alpha, \beta)$ این است که دو رابطه‌ی α و β از طریق پیوند X به هم متصل می‌شوند و A_i را مجموعه‌ی پیوندهای تعریف شده، تاکنون قرار دهیم، هر مرحله می‌تواند، به‌صورت $A_{i+1} = A_i \cup \{X(\alpha, \beta)\}$ نشان داده شود. منظور از الگو، زوجی چون (B, f) است که در آن $B \subseteq A$ و f تابعی است که نحوه‌ی حضور هر عضو از B را در تقطیع جمله نشان می‌دهد. مبنای استخراج الگوهای اولیه [۲۴] می‌باشد و در ادامه، دیگر الگوها از [۲۵] استخراج شده است. در بازنمایی دستگاه مطابقت و نیز گروه‌های فعلی از [۲۶] استفاده شده است. هم‌چنین نگاشت حروف فارسی به لاتین مطابق [27] می‌باشد.

۴-۱- {S (فاعل، فعل)}

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

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



اگر مفعول با - ی همراه باشد، پیش و پس از آن
 [جمله‌ی (۹) شکل ۵ (ب) و جمله‌ی (۱۰) شکل ۵ (پ)].
 اگر بدون نشانه باشد پس از متمم [جمله (۱۱) و شکل
 ۵ (ت)].

ماکروی <verb-t> برای اصلاح تعریف افعال متعدی،
 بازنویسی می‌شود.

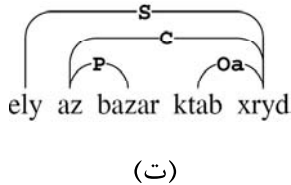
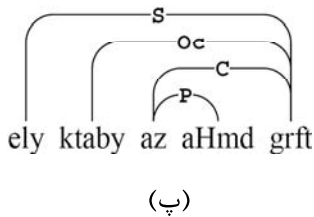
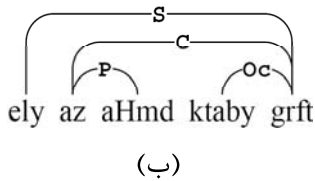
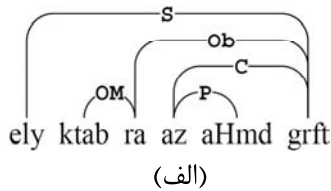
<verb-t> : (({C-} & Ob-) or (Oa- & {C-}) or (Oc-
 & {C-}) or ({C-} & Oc-)) & S-;

(۸) علی کتاب را از احمد گرفت.

(۹) علی از احمد کتابی گرفت.

(۱۰) علی کتابی از احمد گرفت.

(۱۱) علی از بازار کتاب خرید.



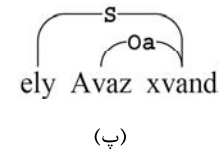
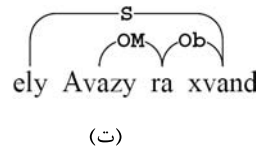
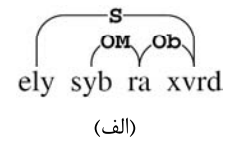
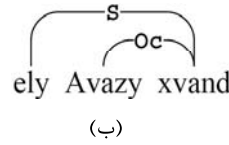
شکل ۵. جایگاه متعدی در جملات با فعل متعدی
 (الف) مفعول با نشانه "را" (ب) و (پ) مفعول با نشانه "ی"
 (ت) مفعول بدون نشانه

(۳) علی سیب را خورد.

(۴) علی آوازی خواند.

(۵) علی آواز خواند.

(۶) علی آوازی را خواند.



شکل ۳ (الف) مفعول با نشانه "را" (ب) مفعول با نشانه "ی"
 (پ) مفعول بدون نشانه (ت) مفعول با هر دو نشانه را و ی

"را" گاهی ممکن است، نشانه‌ی مفعول نباشد؛ این
 موارد در بخش ۴-۸ بررسی شده است.

۴-۴- (متتم، حرف اضافه P، فعل، حرف
 اضافه C)

متتم، اسم یا گروهی اسمی است که بعد از حرف
 اضافه می‌آید و معنی فعل را کامل می‌کند. در جملاتی که
 فعل آن‌ها لازم است، متمم قبل از فعل قرار می‌گیرد [جمله
 (۷) و شکل ۴].

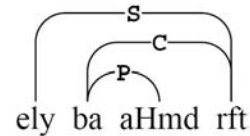
(۷) علی با احمد رفت.

نشانه‌ی 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+;

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



شکل ۴. متمم برای افعال لازم

متتم در جملاتی که مفعول دارند، نسبت به آن
 جایگاه‌های متفاوتی می‌تواند اتخاذ کند [۲۵].

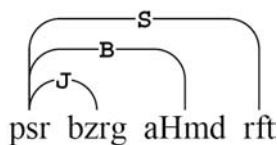
اگر مفعول نشانه‌ی "را" داشته باشد، قبل از متمم می‌آید

[جمله‌ی (۸) و شکل ۵ (الف)].

۴-۵- {T (فعل، تمییز)}

تعریف صفات، شباهت زیادی به اسامی دارد و دلیل آن این است که صفات می‌توانند، جایگاه اسم را اتخاذ کنند. نکته‌ی مهمی که در این رابطه باید تذکر داده شود، وجود براکت اطراف @J+ است. در برخی مراجع، فرض می‌شود که عناصر درون حوزه‌ی اضافه نمی‌توانند توصیف‌های دیگری بپذیرند؛ اما موارد زیادی وجود دارد که این فرض را نقض می‌کند [۲۹]. این مسأله باعث ابهام‌های فراوانی می‌شود که جز با تحلیل معنایی قابل رفع نمی‌باشد. در سیستم کنونی با اعمال هزینه بر روی پیوند @B+ و @J+ سعی می‌شود برتری به پیوندبندی‌هایی داده شود که در آن اسم در جایگاه توصیف‌گر، خود توصیف‌گر دیگری نمی‌پذیرد؛ هرچند که این قاعده، گاهی اوقات به پیوندبندی‌های نامعتبری می‌انجامد. هر براکت، معادل افزایش هزینه به میزان ۱ واحد است. باز در این رابطه، منظور از () عدم تشکیل پیوند است. به طور خلاصه منظور از رابطه‌ی بالا این است که هزینه‌ی تشکیل پیوند B یا J در جایگاه توصیف‌گر، ۱ است. برای اطلاع بیشتر از جزئیات سیستم هزینه‌گذاری گرامر پیوندی، می‌توان به [۱۷] رجوع کرد.

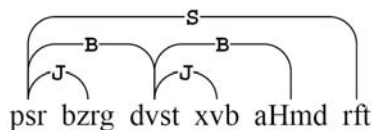
(۱۳) پسر بزرگ احمد رفت.



شکل ۷. وابسته‌های اسم [جمله (۱۳)]

باید توجه داشت که گرامر پیوندی، بدون توجه به ابهام‌های گوناگون، به‌خوبی خاصیت بازگشتی بودن زبان را پوشش می‌دهد؛ برای مثال در جمله‌ی (۱۴) 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 [۲۶].

(۱۴) پسر بزرگ دوست خوب احمد رفت.

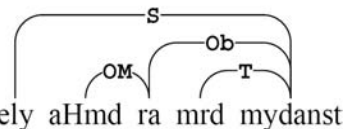


شکل ۸. بازگشتی بودن وابسته‌های اسم و مدل‌سازی آن‌ها به کمک گرامر پیوندی

تمییز با برخی افعال به‌کار می‌رود (مثل دانستن) و معنی آن را کامل می‌کند که از آن به‌عنوان نوعی متمم یا متمم مسندی مفعول [۲۸] یاد می‌شود. جایگاه آن قبل از فعل و بعد از مفعول با نشانه‌ی "/" است و از طریق پیوند T به فعل متصل می‌شود. از ماکروری <verb-ic> برای تعریف این افعال استفاده می‌کنیم، مثلاً در جمله‌ی (۱۲) مرد، نقش تمییزی دارد (شکل ۶)

(۱۲) علی احمد را مرد می‌دانست.

- <noun> : S+ or M+ or Oa+ or OM+ or P- or T+;
- <verb-ic> : ({T-} & {C-} & Ob-) or (Oa- & C-) or (Oc- & {C-}) or ({C-} & Oc-) & S-;



شکل ۶. نقش تمییزی برای افعال ناقص [جمله (۱۲)]

۴-۶- وابسته‌های اسم: {J (موصوف)، J ،

{B (مضاف الیه، مضاف)}

این وابسته‌ها شامل صفت و مضاف الیه هستند. صفت از طریق پیوند J و مضاف‌الیه از طریق پیوند B به اسم، متصل می‌شود [جمله‌ی (۱۳) و شکل ۷]. برای سادگی پیوندهای اسم را در دو دسته قرار می‌دهیم: پیوندهایی که در آن اسم وابسته به فعل باشد (فاعل، مفعول و تمییز) که به کمک ماکروری <noun-vc> خلاصه می‌شوند و دیگری (مضاف الیه و متمم) که ماکروری <noun-nc> به‌همین منظور به‌کار می‌رود. اعمال مشابهی هم برای صفات صورت می‌گیرد.

- <noun-nc> : (B- & (([@J+] or ())) & ([@B+] or ())) 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> : (J- & (([@J+] or ())) & ([@B+] or ())) or (P- & ({@J+} & {@B+}));
- <adj-vc> : {@J+} & {@B+} & (S+ or M+ or Oa+ or OM+ or T+);
- <adj> : <adj-jc> or <adj-vc>;



(حتماً) و ترتیب (دوبه‌دو) می‌باشد که از طریق پیوند Ab به فعل متصل می‌شوند.
(۱۸) علی پیاده آمد.

<adv-b>: Ab+;

۴-۷-۳- قیدهایی که پیش و پس از فاعل قرار می‌گیرند [جملات (۱۹) و (۲۰) 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 اتصال گر A قابلیت اتصال با هر دو نوع Aa و Ab را دارد. (ر.ک. قاعده‌ی تطابق بخش ۳)

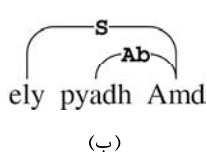
(۱۹) دیروز علی کتاب خرید.

(۲۰) علی دیروز کتاب خرید.

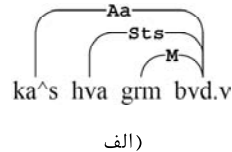
<adv-c>: A+;

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

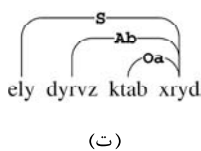
- <modal> : M- & (@Ab-) & S- & (@Aa-);
 <verb-it> : {C-} & (@Ab-) & S- & (@Aa-);
 <verb-ic> : (({T-} & {C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & (@Ab-) & S- & (@Aa-);
 <verb-t> : (({C-} & Ob-) or (Oa- & {C-}) or (Oc- & {C-}) or ({C-} & Oc-)) & (@Ab-) & S- & (@Aa-);



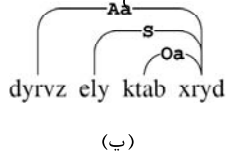
(ب)



(الف)



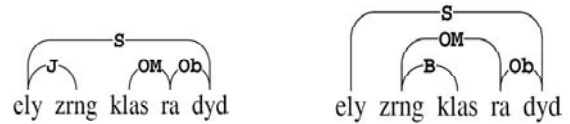
(ت)



(پ)

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

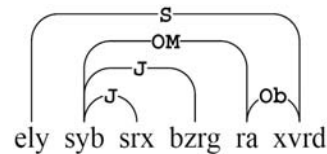
البته در ساختارهای پیچیده‌تر، ابهام زیادی رخ می‌دهد؛ مثلاً در جمله‌ی (۱۵) دو تقطیع برای جمله وجود دارد که در شکل ۹ مشخص است.
(۱۵) علی زنگ کلاس را دید.



شکل ۹. ابهام در هنگام متوالی گشتن اسامی در نقش مضافی و صفت [جمله (۱۵)]

وجود @J- برای این است که یک اسم ممکن است هم‌زمان چندین صفت را اتخاذ کند [جمله‌ی (۱۶) و شکل ۱۰].

(۱۶) علی سیب سرخ بزرگ را خورد.



شکل ۱۰. صفات متوالی برای یک اسم [جمله (۱۶)]

۴-۷-۷- {فعل، قید} A

قید جایگاه ثابتی در جمله ندارد؛ اما با توجه به نوع آن، می‌توان جایگاهی را که معمولاً اتخاذ می‌کند پیش‌بینی کرد. از این حیث، قیود را می‌توان به سه دسته تقسیم کرد.
(۲۵)

۴-۷-۱- قیدهایی که پیش از فاعل قرار می‌گیرند [جمله‌ی (۱۷) و شکل ۱۱ (الف)].

این دسته، شامل قیدهایی تمنا (کاش)، تأسف (افسوس) و تحسین و مسرت (خوشا) می‌باشد که از طریق پیوند Aa به فعل متصل می‌شوند.
(۱۷) کاش هوا گرم بود!

<adv-a>: Aa+;

۴-۷-۲- قیدهایی که پس از فاعل قرار می‌گیرند [جمله‌ی (۱۸) و شکل ۱۱ (ب)].

این دسته، شامل قیدهایی مکان (هرجا)، کمیت (کمی)، کیفیت (زیبا)، حالت (پیاده)، پرسش (کجا)، تصدیق

سایر صورت‌ها به شکل مشابهی انجام می‌گیرد. با در نظر گرفتن تنوع‌های مختلف در ساختار فعل، از ساخت ماضی مطلق یک فعل مثل رفت، فعل‌های می‌توان رفت، می‌شد رفت، بشود رفت، خواهد رفت و رفت ساخته می‌شود. هریک از اعضای رابطه‌ی (۱) را می‌توان به صورت زیر تعریف کرد:

۴-۹-۱- {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 به صورت 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 دهنده‌ی زمان‌های حال، گذشته و التزامی هستند.

ب) فعل/دشتن که قابل صرف برای صیغه‌های مختلف می‌باشد.

da[^]st : GBP+;

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

xvahd : GCP+;

۴-۹-۲- {(فعل، عنصر غیر فعلی) K}

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

baz va ps ... : K+;

۴-۹-۳- {(، عنصر ساختمانی حالت، فعل) RT}

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

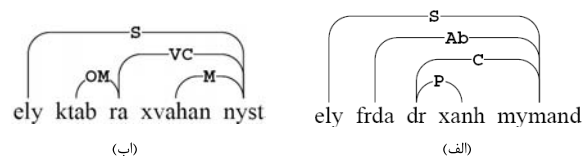
bvd : RT+;

۴-۸- کاربردهای جانبی "را"

علی‌رغم این‌که کاربرد "را" در بیشتر موارد نشانه‌ی مفعولی است، اما حالاتی وجود دارد که "را" نقشی متفاوت بر عهده دارد که از آن نمونه می‌توان به وضعیتی اشاره کرد که همراه با گروه اسمی، نقش قید زمان یا مکان پیدا می‌کند [پیوند AC با فعل] (جمله (۲۱)) و یا نقش متمم فعلی به خود می‌گیرد. [پیوند VC با فعل]، (جمله (۲۲)) [۲۸]. برای پوشش این حالات باید تعاریف ra، افعال ربطی، قیود و افعالی چون، ماندن به صورت زیر تغییر کند.
(۲۱) علی فردا را در خانه می‌ماند.
(۲۲) علی کتاب را خواهان نیست.

ra : OM- & (Ob+ or OC+ or AC+);
<modal> : M- & {@Ab-} & {OC-} & {@Ab-} & S- & {@Aa-};
<adv-c> : A+ or OM+;
mymand : {C-} & {AC-} & {@Ab-} & S- & {@Aa-};

نتیجه‌ی تقطیع جملات بالا در شکل ۱۲ آمده است.



شکل ۱۲. کاربردهای جانبی را (الف) نشانه‌ی قید، (ب) نشانه‌ی متمم فعلی

۴-۹-۴- گوناگونی های گروه فعلی

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

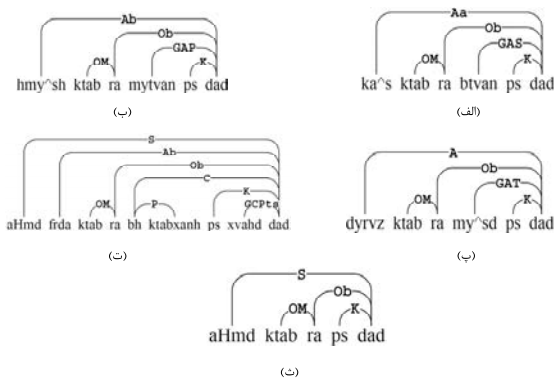
فعل ناقص + عنصر غیر فعلی + فعل واژگانی + عنصر سازنده‌ی مجهول + عنصر سازنده‌ی حالت

با در نظر گرفتن فعل واژگانی به عنوان هسته‌ی این ساختار، می‌توان همه‌ی دستگاه‌های هم‌زمان گروه فعلی را ساخت. فعل واژگانی می‌تواند از یکی از ده صورت مختلف فعل انتخاب شود، برای سادگی ما تنها صورت ماضی مطلق (در ساخت‌های معلوم و مجهول) را در نظر می‌گیریم، تعریف

- & <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 پیوندی به کار رفته در تعاریف بالا را می توان از طریق مثال های زیر نمایش داد (شکل ۱۳).

- (۲۳) کاش کتاب را بتوان پس داد!
 (۲۴) همیشه کتاب را می توان پس داد.
 (۲۵) دیروز کتاب را می شد پس داد.
 (۲۶) احمد فردا کتاب را به کتابخانه پس خواهد داد.
 (۲۷) احمد کتاب را پس داد.



شکل ۱۳. پیوندی حاصل برای جملات (۲۳) تا (۲۷) | به ترتیب الف) تا (ث)

۵- ساخت مجهولی: {سازندهی مجهول، صفت مفعولی}L

در ساخت مجهولی، همه ی تغییرات لازم در زمان های مختلف باید به عنصر سازندهی مجهول منتقل شود. برای مثال همان ساخت ماضی مطلق فعل شدن را در نظر می گیریم. عنصر سازندهی مجهول (فعل شدن) همانند سایر افعال تعریف می شود با این تفاوت که یک پیوند L با اسم مفعول تشکیل می دهد و فاقد مفعول می باشند. ملزومات پیوندی افعال متعدی و نیازمند تمییز را در حالت های 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 مفعولی هایی که به فعل شد، متصل شده اند (متعدی یا نیازمند تمییز)، به ترتیب از دو نوع پیوند متفاوت Lt و Lic برای اتصال به فعل شد

۴-۹-۴- دستگاه های مختلف فعلی

همان طور که ذکر شد، در این بخش صیغه های مختلفی که از ساخت ماضی مطلق ساخته می شود، در نظر می گیریم. ساخت سایر صیغه ها به شکل مشابهی صورت می گیرد. برای هر یک از این صیغه ها، ماکرویی تعریف می شود که در بردارندهی تعریف آن صیغه می باشد.

الف) التزامی غیر شخص (جمله ۲۳)

- <ISUB> : {K-} & GAS- ;
 ب) گذشته اخباری غیر شخصی [جملات (۲۴) و (۲۵)]
 <IPIN> : {K-} & (GAP- or GAT-);

در فرم های غیر شخصی جایگاه فاعل، خالی می ماند؛ در نتیجه، تعاریف قبلی (<modal> و غیره) باید اصلاح شود و اتصال گر S حذف شود. هم چنین با توجه به عدم وجود فاعل، اتصال گر قید به A تغییر می یابد. چون در این حالت بین انواع قیدی که در بخش ۴-۷ 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 ۲۶]

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

ت) گذشتهی غیر استمراری مطلق [جمله ۲۷]

- <NCSP> : {K-};

و در نهایت ماکرویی برای تعریف افعال ماضی مطلق ماکروی زیر مانند قبل به ترتیب برای تعریف ساخت ماضی مطلق افعال اسنادی، لازم، نیازمند تمییز و متعدی به کار می رود.

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

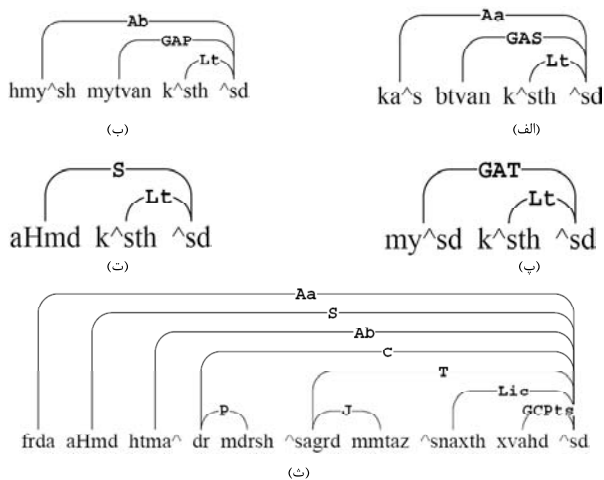
(۳۱) احمد کشته شد.

(۳۲) فردا احمد حتماً در مدرسه شاگرد ممتاز شناخته

خواهد شد.

۶- مطابقتی فعل و فاعل

در این بخش و بخش بعد، تعریف اسامی و افعال را اندکی پیچیده تر می‌کنیم تا نحوه‌ی ساخت انواع فعل در زبان فارسی و همچنین نحوه‌ی 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 به‌سادگی قادر به پشتیبانی آن می‌باشد. برای پوشش دادن دستگاه مطابقتی زبان فارسی، نیاز است ماکروبی که برای تعریف اسامی در نظر گرفته بودیم یعنی <noun> (در اصل <noun-vc>)، بسط داده شود، چرا که تطابق در زبان فارسی برای اسامی بی‌جان، 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 بالا را بررسی می‌کنیم.



شکل ۱۴. پیوندبندی حاصل برای جملات (۲۸) تا (۳۲) [به ترتیب (الف) تا (ث)].

۶-۱- فاعل بی‌جان

در این حالت، فاعل مفرد باید با فعل مطابق باشد، اما در حالت جمع، در تعداد فاعل اجبار وجود ندارد [جملات (۳۳) و (۳۴)]. دو ماکروبی <noun-in-s> و <noun-in-p>

استفاده می‌کنیم. از دو ماکروبی <PP-t> و <PP-ic> به ترتیب برای تعریف این دو نوع صفت مفعولی استفاده می‌کنیم. دو ماکروبی <SP-t-p> و <SP-ic-p> نیز مانند حالت معلوم، برای تعریف انواع ساخت‌ها و زمان‌های فعل مجهولی که از اتصال دو نوع صفت مفعولی مذکور به فعل شده، ساخته می‌شود، به کار می‌رود. تنها تفاوت، این است که این ماکروها قبلاً در تعریف ساخت ماضی مطلق فعل به کار می‌رفتند و در اینجا در تعریف ساخت ماضی مطلق عنصر سازنده مجهول به کار می‌رود. همچنین باز باید توجه کرد که ماکروبی <FUT> در اینجا قابل استفاده نیست؛ چرا که 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 به صورت دو ماکروبی <FUT-t-p> و <FUT-ic-p> بازنویسی شده است که t و ic مانند قبل به نوع فعل دلالت می‌کند.

<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 به عنوان یک فعل واژگانی مستقل به کار رود، به همین دلیل می‌توان ماکروبی <SP-modal> را به تعریف آن اضافه کرد. این کار می‌تواند در مدخل دیگری از واژه‌نامه انجام شود.

مثال‌های (۲۸) تا (۳۲) می‌تواند در توجیه تعاریف بالا مفید باشد. بالاخص مثال (۳۲) نمونه‌ای از یک جمله‌ی نسبتاً پیچیده برای یک فعل نیازمند تمییز است (شکل ۱۴).

(۲۸) کاش بتوان کشته شد!

(۲۹) همیشه می‌توان کشته شد.

(۳۰) می‌شد کشته شد.

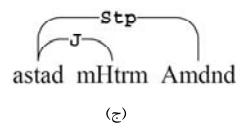
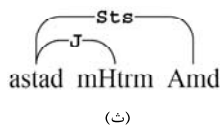
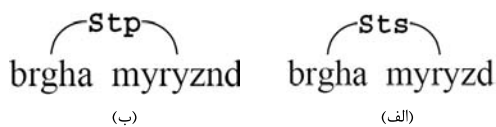
- <noun-vc-hu-fs> : { @J+ } & { @B+ } & (Sfs+ or M+ or Oa+ or OM+ or T+);
 <noun-vc-hu-ts> : { @J+ } & { @B+ } & (St+ or M+ or Oa+ or OM+ or T+);
 <noun-hu-ts> : <noun-nc> or <noun-vc-hu-fs>;
 <noun-hu-tp> : <noun-nc> or <noun-vc-hu-ts>;

۴-۶- 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 ۴-۹-۳) و عنصر حالت (بخش ۴-۹-۳) به شکل مشابهی اصلاح می‌شوند (رج ضمیمه‌ی الف).

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

در نهایت می‌توان صحت تقطیع قواعد را برای جملات (۳۳) تا (۳۸) سنجید (شکل ۱۵).



شکل ۱۵. تطبیق فاعل و فعل برای جملات (۳۳) تا (۳۸) به ترتیب (الف) تا (ج)

برای تعریف این اسامی به ترتیب در حالت‌های مفرد و جمع به کار می‌رود.

(۳۳) برگ‌ها می‌ریزند.

(۳۴) برگ‌ها می‌ریزد.

- <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>;

همان‌طور که مشخص است، تنها پیوندی که تغییر کرده، پیوند مابین فاعل و فعل است (S)، زیرنویس‌های t و s برای نشان دادن سوم شخص و مفرد بودن به کار می‌روند. همان‌گونه که مشاهده می‌شود، مطابقت در گرامر پیوندی عموماً از طریق زیرنویس‌ها انجام می‌گیرد؛ دلیل این کار با توجه به نحوی تطابق پیوندها که در بخش ۳ ذکر شد، منطقی به نظر می‌رسد.

۲-۶- فاعل جاندار غیر انسان

در این حالت، بین فاعل و فعل تطابق وجود دارد [جملات (۳۵) و (۳۶)]. <noun-an-s> و <noun-an-p> برای تعریف این اسامی به ترتیب در حالت‌های مفرد و جمع به کار می‌رود.

(۳۵) گربه آمد.

(۳۶) گربه‌ها آمدند.

- <noun-vc-an-s> : { @J+ } & { @B+ } & (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>;

۳-۶- مسندالیه انسان

در همه‌ی حالات، مطابقت وجود دارد، جز در حالتی که مسندالیه سوم شخص مفرد باشد، در این حالت در تعداد مطابقت ضروری نیست [جملات (۳۷) و (۳۸)]. 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:

(۳۷) استاد محترم آمد.

(۳۸) استاد محترم آمدند.

۷- نتیجه‌گیری و کارهای آینده

به‌روشنی، از مثال‌ها و مراحل تکمیل گرامر، به‌نظر می‌رسد که توسعه‌ی این گرامر به‌سادگی صورت می‌گیرد. آنچه در این مقاله بررسی شد، نحوه‌ی تعریف واژه‌نامه برای تحت پوشش قرار دادن نقش‌های اساسی و الگوهای اولیه برای بندهای معین 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 (هسته‌ای یا خوشه‌ای) اشاره کرد. البته ایجاد بسترهای آزمایش مناسب و استفاده از این گرامرها، جهت ادغام در سیستم‌های دیگر، از اهداف آینده‌ی این پروژه می‌باشد.

۸- خلاصه

گرامر پیوندی به‌عنوان صورت‌بندی‌ای برای بازنمایی زبان‌های مستقل از متن، به علت لغوی بودن و در نتیجه کاهش پیچیدگی دستور زبان و توزیع آن در سطح همه‌ی کلمات و نیز مقیاس‌پذیری آن، می‌تواند به‌عنوان بدیلی برای گرامرهای با ساخت گروهی به‌کار رود. این گرامر بدون نیاز به مفهوم واحد و تجزیه‌ی مستقیم، همه‌ی خواص یک گرامر برای بازنمایی دانش زبان طبیعی را دارد. 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 مجموعه‌ای از ملزومات پیوندی را فراهم آوردیم که قادر به پردازش جملات ساده و ابتدایی زبان باشد. تطبیق فعل و فاعل در زبان فارسی، نمونه‌ای از خواص لغوی بود که گرامر پیوندی با توجه به لغوی بودن، به‌سادگی قادر به اداری آن‌ها می‌باشد. در هر مرحله با افزودن مثال‌هایی، سعی کردیم روش خود را ملموس‌تر نماییم. با آزمایش مرحله به مرحله‌ی قواعد، به کمک تقطیع‌گر گرامر پیوندی از مجموعه‌ی به‌دست آمده، یک مجموعه‌ی سازگار استخراج شد (کل واژه‌نامه در ضمیمه‌ی الف گردآوری شده است).

۹- مراجع

- [۱] عبدالله زاده، ا.، "یادداشت‌های درس پردازش زبان طبیعی"، دانشگاه صنعتی امیرکبیر، ۱۳۸۳
- [2] Chomsky, N. "Syntactic Structures", The Hague: Mouton.
- [3] Naur, P., "Revised Report on the Algorithmic Language, ALGOL 60", Communications of the ACM 6(1). January 1958.
- [۴] دبیرمقدم، م.، "زبان‌شناسی نظری: پیدایش و تکوین دستور زایشی"، ویرایش ۲، سازمان مطالعه و تدوین کتب علوم انسانی و دانشگاهها (سمت)، ۱۳۸۳.
- [5] Gazdar, Gerald and G. Pullum. "Generalized Phrase Structure Grammar: A Theoretical Synopsis", Indiana University Linguistics Club, Bloomington, Indiana. 1982.
- [6] Pollard, C., and Sag I. A., "Head-driven Phrase Structure Grammar", Chicago Univ Press, 1994
- [7] Russel, S. and Norvig, P., "Artificial Intelligence: A Modern Approach", Second Edition, Englewood Cliffs, NJ: Prentice-Hall, 2003
- [8] Kaplan, R. and Bresnan, J., "Lexical Functional Grammar: a formal system for Grammatical representation", In Joan Bresnan (ed.), The Mental Representation of Grammatical Relations. MIT Press, Cambridge, Mass, 173-281, 1982.
- [9] Joshi, A. K. and Schabes. Y., "Tree-adjoining grammars and lexicalized grammars", In Tree Automata and LGS. Elsevier Science, Amsterdam, 1992
- [10] Meřuk, I. A. "Dependency Syntax: Theory and Practice", State University of New York Press 1988.
- [۱۱] فیلی، ه. و قاسم‌ثانی، غ.، "استفاده از گرامر درخت-افزایشی برای ترجمه انگلیسی به فارسی"، مجموعه مقالات نهمین کنفرانس سالانه انجمن کامپیوتر ایران، جلد اول ص ۶۳۹-۶۴۷، ۱۳۸۲

[۱۲] فیلی، ه. و قاسم‌ثانی، غ.، "استفاده از گرامر یکسان ساز برای ترجمه انگلیسی به فارسی"، مجموعه

[۲۴] خانلری. پ.، "دستور زبان فارسی"، ویرایش دوم، ۱۳۸۲، انتشارات توس.

[۲۵] انوری، ح.، گیوی، "حسن، دستور زبان فارسی ۲"، ویرایش ۲، چاپ بیست و سوم، ۱۳۸۳، انتشارات فاطمی

[۲۶] باطنی. م.، "توصیف ساختمان دستوری زبان فارسی"، ۱۳۸۳، انتشارات امیرکبیر.

[27] Megerdooian, K., "Unification-based Persian morphology", In proceedings of CICLing , Alexander Gelbukh, Center of Investigation on Computation-IPN, Mexico, 2000.

[۲۸] مشکوة الدینی، م.، "دستور زبان فارسی بر اساس نظریه گشتاری"، ویرایش دوم، ۱۳۷۹، دانشگاه فردوسی مشهد.

[۲۹] معظمی، آ. و درزی، ع.، "ساخت حوزه اضافه در گروه اسمی زبان فارسی"، دومین کارگاه پژوهشی زبان فارسی و رایانه، دانشگاه تهران، تیرماه ۱۳۸۵.

مقالات دهمین کنفرانس سالانه انجمن کامپیوتر ایران، جلد اول، ص ۹۳۶-۹۴۴، ۱۳۸۳

[۱۳] آیت، م.، "یک گرامر محاسباتی برای زبان فارسی"، پایان‌نامه‌ی کارشناسی ارشد رشته‌ی مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، ۱۳۸۰

[14] Radford, A., "Transformational Syntax", Cambridge University Press, 1981.

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

[16] Sleator, D. and Temperley, D., "Parsing English with a Link Grammar", Third International Workshop on Parsing Technologies, 1993

[17] Sleator, D. and Temperley, D., "Parsing English with a Link Grammar", Carnegie Mellon University Computer Science technical report CMU-CS-91-196, October 1991

[18] Lafferty, J., Sleator, D. and Temperley, D., "Grammatical Trigrams: A Probabilistic Model of Link Grammar", Proceedings of the AAAI onference on Probabilistic Approaches to Natural Language, October, 1992.

[19] Madhyastha, H.V., Balakrishnan, N., Ramakrishnan, K.R., "Event Information Extraction Using Link Grammar", 13th International WorkShop on Research Issues in Data Engineering: Multi-lingual Information Management (RIDE'03), October 2004.

[20] 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 (ICTAI'03). 2003

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

[22] Carol, L., "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

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

ضمیمه‌ی (الف)

<noun-nc>	: (B- & (((@J+) or ()) & ([@B+] or ()))) 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+);
<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>	: (J- & (((@J+) or ())) & ([@B+] or ()))) or (P- & ({@J+} & {@B+}));
<adj-vc>	: {@J+} & {@B+} & (S+ or M+ or Oa+ or OM+ or T+);
<adj>	: <adj-jc> or <adj-vc>;
<prep>	: P+ & C+;
ra	: OM- & (Ob+ or OC+ or AC+);
<adv-a>	: Aa+;
<adv-b>	: Ab+;
<adv-c>	: A+ or OM+;
<modal-ts>	: M- & {@Ab-} & {OC-} & {@Ab-} & Sts- & {@Aa-};
<modal-tp>	: M- & {@Ab-} & {OC-} & {@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+;
xvahd	: 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>	: Lt+;
<PP-ic>	: Lic+;
<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-};
<SP-t-ts-p>	: GCPTs- & Lic- & {K-};
<SP-ic-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	: <SP-t-ts-p> or <SP-ic-ts-p>;
hva Avaz syb psr ktab bazar klas	
brg ktabxanh mdrsh xanh	: <noun-in-s>;
brgha	: <noun-in-p>;
grbh	: <noun-an-s>;
grbhha	: <noun-an-p>;
^sagrd aHmd karmnd ely mrd dvst	
astad	: <noun-hu-ts>;
ktaby Avazy	: <noun-indef>;
dr az ba bh	: <prep>;
mmtaz zrng srx bzrg xvb grm	
xvahan mHtrm	: <adj>;
ka^s	: <adv-a>;
htma^ hmy^sh pyadh	: <adv-b>;
dyrvz frda	: <adv-c>;
my^svd.v bvd.v ast nyst	: <modal-ts>;
my^svnd	: <modal-tp>;
mydanst	: <verb-ic-ts>;
myryzd	: <verb-it-ts>;
mymand	: {C-} & {AC-} & ({@Ab-} & Sts- & {@Aa-});
Amdnd myryznd	: <verb-it-tp>;
dad grft xvrd xvand xryd dyd k^st	: <SP-t-ts>;
Amd rft	: <SP-it-ts>;
^snaxt	: <SP-ic-ts>;
k^sth	: <PP-t>;
^snaxth	: <PP-ic>;



آرمین سجادی در سال ۱۳۶۰ در شهر سنندج متولد شد. تحصیلات تا مقطع دیپلم را در شهر سنندج سپری و دیپلم متوسطه‌ی خود را در سال ۱۳۷۸ دریافت کرد. وی

تحصیلات خود در مقطع کارشناسی را در رشته‌ی مهندسی کامپیوتر در دانشگاه علم و صنعت ایران (سال ۱۳۸۳) و کارشناسی ارشد را در رشته‌ی هوش مصنوعی از دانشگاه صنعتی امیر کبیر (پلی تکنیک تهران) (سال ۱۳۸۵) به پایان رسانید. زمینه‌ی تخصصی مورد علاقه‌ی ایشان پردازش نمادی و زبان‌شناسی محاسباتی می‌باشد.

نشانی (رایانامک) پست الکترونیکی ایشان عبارت است از:

asajadi@aut.ac.ir



احمد عبدالله زاده کارشناسی خود را

در رشته‌ی حسابداری از دانشگاه تهران (سال ۱۳۵۴) و کارشناسی ارشد را در رشته‌ی علوم کامپیوتر از دانشگاه وست کوست لس آنجلس در آمریکا در سال ۱۳۵۹ دریافت کرد. وی همچنین در

سال ۱۳۶۹ موفق به اخذ دکترای تخصصی در رشته‌ی علوم کامپیوتر از دانشگاه بریستول لندن شد.

ایشان به عنوان استاد مدعو در سال‌های ۱۳۷۹ تا ۱۳۸۱ در دانشگاه‌های مریلند آمریکا و ارسای پاریس حضور داشته است. وی هم‌اکنون دانشیار دانشکده‌ی مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی امیرکبیر می‌باشد.

زمینه‌های تخصصی ایشان عبارتند از: بازیابی اطلاعات، تکنیک‌های هوش مصنوعی، سیستم‌های خبره، پردازش زبان طبیعی، سیستم‌های تصمیم‌یار، بازنمایی دانش و مهندسی نرم‌افزار.

نشانی (رایانامک) پست الکترونیکی ایشان عبارت است از:

ahmad@aut.ac.ir