



## Agglutinative-Analytic Morphology of Persian: A Distributed Morphology Approach

Shabnam Majidi<sup>1</sup> & Mahinnaz Mirdehghan Farashah<sup>2</sup>

(147-172)

### Abstract

The present research is aimed to analyse the morphological typology of Persian on the basis of the framework of Distributed Morphology (DM). According to the findings of DM, the predominant tendency of Persian's morphological type is defined here on a two-dimensional chart; a chart provided by Aikhenvald (2007) from the conjugation typology of languages, which itself was defined on the basis of the criteria proposed by Sapir. The horizontal axis in this diagram deals with the degree of internal complexity of words and its horizontal axis includes the transparency of morphological boundaries between the morphemes within a grammatical word. Accordingly, the main research questions are to determine: the position of Persian language on the basis of "degree of internal complexity of words", as well as, its typology on the basis of "border transparency criteria". To analyse the grammatical structures of Persian, a variety of diverse structures, among different Persian text types, including stories, textbooks, scenarios, and research papers was gathered, which contained a total number of nine thousand and six hundred words. The investigation on the first question on the basis of Greenberg's quantitative standards illustrates that, by achieving a number of one thousand and five hundred and six words from the above-mentioned index, the orientation of today's Persian language, according to the "degree of internal complexity of words" or what is called as "word combinability criteria", is towards "analytic" language type. Considering the second question, the research demonstrates that in view of the "border transparency criteria" for morphemes which indicates ways to connect morphemes, Persian language is more inclined towards "agglutinating" language type. Finally, it can be concluded that based on the findings of DM, the dominant tendency of morphosyntactic structures of Persian is towards "agglutinating-analytic" languages. Therefore, the research findings in considering Persian within the DM framework approves its grouping within agglutinating-analytic language types, though in comparison to the previous studies, the reduction of the statistical results here illustrates the main tendency of Persian's morphological type towards analytic languages and its movement towards becoming more analytic.

**Keywords:** distributed morphology, morphological typology, quantitative approach, analytic language, agglutinating languages.

Received: June, 20, 2020; Accepted: December, 5, 2020

doi  
10.22059/jolr.2020.304807.666611  
Print ISSN: 1026-2288-Online ISSN: 2676-3362  
<https://jolr.ut.ac.ir>

1. Ph.D. Candidate of Linguistics, Shahid Beheshti University, Tehran, Iran.

2. Corresponding Author Email: M\_Mirdehghan@sbu.ac.ir

Associate Professor Linguistics, Shahid Beheshti University, Tehran, Iran.

### **1. Introduction**

In contrast to lexical approaches, a distributed morphology approach produces structures using formal features and seeks to provide the same analysis of the construction of composite structures (word, group, and sentence). This theoretical approach claims that analytic, synthetic, and polysynthetic structures are produced in all isolating, agglutinating, and fusional languages in the syntax field. Thus, the present study seeks to develop a new analysis of determining the morphological typology of Persian language in the context of distributive morphology, by focusing on the following two questions:

Question 1: Where is the position of the Persian language on the axis of complexity within grammatical structures?

Question 2: What is the Persian's morphological type from the viewpoint of "border transparency criteria"?

### **2. Theoretical foundations**

This section provides a brief overview of the morphological types of languages and the criteria for typing languages based on the word production process and the formation process of grammatical structures. Accordingly, two criteria are introduced as follows: The first criterion is border transparency criteria – Sapir (1921) – which recognizes three isolating, agglutinating, and fusional types. The second criterion is the degree of internal complexity of words, in which three types of isolating, agglutinating, and fusional languages can be identified. This section also refers to the generalities of distributed morphology, along with an analysis of descriptive and derivative morphemes, which is different from the traditional morphology approach.

### **3. Background**

The current section elaborates on the typology research and studies related to this article. Therefore, the research of Aikhenvald (2007), Samareh (1990), Sharifi and Akhlaghi (2013), and Dabir Moghaddam (2013) are reviewed. Then, the analysis and results of Anoushe (2018), are discussed to identify morpheme segmentation based on the distributed morphology approach.

### **4. Study and analysis of morphological data**

In the first section, a variety of linguistic body of Persian texts is collected. The selected texts do not belong to a specific style, context, field, and specialty, and contain various styles such as writing style, conversational style, and scientific and fictional styles from different fields, including social, cultural, religious, political, sports, economics, medicine, and health, engineering sciences, language and literature, architecture, history, industry, and tourism. Then, the structure of the Persian language from a DM approach is examined and analyzed based on the findings of this grammatical concept based on Sapir criteria and Greenberg's (1954) quantitative approach, on the basis of the two-dimensional Aikhenvald axis (2007).

The morpheme segmentation of body words from a non-lexical DM shows manifestations of the morphosyntactic structures of the Persian language: in Persian, the "future structure" is made analytically with the auxiliary verb "[xastan]: will"; That is, the grammatical meaning is expressed not through affixation, but through independent words. The current form of the past in the

Persian language is made synthetically. The simple past in Persian is formed by adding the word “d/t” affix to the present root; like the simple past tense in the words “[xordam]: I ate” and “[barnemigardanam]: I did not return.” Structures such as “[hamiše az xoda xaste]: he/she always asked God” and “[az xod razi]: self-satisfied” are examples of fusional or multi-synthetic languages in Persian that have a lot of complexity within their structure. These structures are made as a result of the process of synthesizing different lexical and grammatical elements and turning them into a single word. Structures such as “[dorost dar kenare divare otāg]: right next to the wall of the room” and “[ba tʃand ketan dar dast ]: with a few books in hand” are also examples of isolating languages in Persian. In this type of structure, there is a one-to-one correspondence between morphemes and words, and each morpheme is equal to one word. Words such as “[daneʃmandhayeʃan]: their scientists” and “[pazuhaʃgaran]: researchers” are examples of the type of synthetic languages in Persian, where each word is composed of several morphemes and the boundary between morphemes is clear. In the morphology of the Persian language, the fusion of verbs and roots has a low frequency; For example, the third person singular verb “[ast]: is” is the result of the fusion of the root “[baʃ]: Bash” and the present tense and the suffix “zero”. In the following, we use Greenberg Quantitative indicator to determine the internal complexity of Persian’s grammatical structures. This index refers to determining the degree of the internal complexity of languages and is obtained by dividing the number of Morpheme in a sample volume by the number of words in that sample. According to Table 1 and the horizontal axis of Figure 2, the quantitative value indicates that the morphosyntactic structures of Persian tend to be analytical languages. Then, Persian words are examined in terms of the border transparency criteria in the vertical axis of Aikhenvald. Generally, 64% of the words are made by synthesizing and affixing words in the Persian language morphology system while 27% of morpheme words belong to the type of isolating languages, in which there is a one-to-one correspondence between morphemes and words. Ultimately, only 9% of morpheme words belong to the group of fusional languages. Therefore, it can be acknowledged that the predominant tendency of the Persian's morphosyntactic type is towards the agglutinative language's type according to the border transparency criteria.

### 5. Result

Based on the findings obtained from the distributed morphological typology, suitable analysis and input were provided to determine the morphosyntactic structure of the Persian language. Contemporary Persian is more inclined to the type of Agglutinative languages in terms of the degree of the internal complexity of words to the type of analytical languages in terms of the border transparency criteria. Thus, the distributed morphological typology places the predominant tendency of the Persian language morphological typology in the type of Agglutinative-Analytic languages.

## ساختواژه تحلیلی پیوندی زبان فارسی بر پایه نظریه صرف توزیعی

شبلم مجیدی

دانشجوی دکتری زبان‌شناسی، دانشگاه شهید بهشتی، تهران، ایران

مهین‌ناز میردهقان فراشاه<sup>۱</sup>

دانشیار زبان‌شناسی، دانشگاه شهید بهشتی، تهران، ایران

تاریخ دریافت مقاله: ۹۹/۳/۳۱؛ تاریخ پذیرش مقاله: ۹۹/۹/۱۵

علمی - پژوهشی

## چکیده

در مقاله حاضر انگاره «صرف توزیعی» به عنوان شالوده نظری با هدف تحلیل رده‌شناسی ساختواژی زبان فارسی به کار بسته شده است. بدین ترتیب، بر پایه یافته‌های نظریه صرف توزیعی، گرایش رده ساختواژی زبان فارسی در نمودار دوبعدی آیکنوالد (۲۰۰۷) که بر اساس معیارهای ساپیر (۱۹۲۱) ارائه شده، مورد تحلیل قرار گرفته است. در این پژوهش در پی آنیم تا بدانیم: جایگاه زبان فارسی بر روی محور «پیچیدگی درون ساخت‌های دستوری» کجاست؟ و از منظر «شفافیت مرز میان تکواژها»، زبان فارسی در چه رده زبانی می‌گنجد؟ در رسیدن به اهداف پژوهش، پیکره زبانی متنوعی از متون فارسی، مشتمل بر داستان، درسنامه، فیلمنامه و مقاله‌های مختلف زبان فارسی گردآوری شده است. پیکره یادشده، ۲۴ نوع متن گوناگون معاصر را در برمی‌گیرد و از هر یک از آنان، ۴۰۰ واژه به صورت تصادفی در قالب جمله نمونه‌گیری شده، که در مجموع ۹۶۰۰ واژه را شامل می‌شود. نتایج حاصل از پژوهش حاضر، همراه با حصول عدد ۱،۵۰۶ بر مبنای تحلیل داده‌ها با شاخص کمی گرینبرگ، نمایانگر آن است که در پاسخ به پرسش اول، گرایش زبان فارسی امروز بر اساس معیار «میزان پیچیدگی درون ساخت‌های دستوری واژه‌ها» یا همان معیار «ترکیب‌پذیری واژه‌ها»، به سوی رده زبان‌های «تحلیلی» است. یافته دیگر پژوهش را در پاسخ به پرسش دوم آن می‌توان چنین برشمرد که زبان فارسی بر اساس معیار «شفافیت مرز تکواژها» که نمایانگر راه‌های اتصال تکواژهاست، به رده زبان‌های «پیوندی» گرایش بیشتری دارد. بدین ترتیب، گرایش غالب ساخت‌های صرفی نحوی زبان فارسی از منظر صرف توزیعی که تفاوت‌های بنیادین در نوع و شمار تکواژها با صرف واژه‌گرا دارد، در رده زبان‌های «تحلیلی پیوندی» قرار می‌گیرد. از این‌رو، انگاره دستوری صرف توزیعی نیز تأییدکننده تحلیلی پیوندی بودن رده ساختواژی زبان فارسی است؛ با این تفاوت که کمتر شدن عدد کمی به دست آمده در این پژوهش، نسبت به یافته‌های پژوهش‌های پیشین، حاکی از آن است که بر پایه تقطیع ساختواژی در صرف توزیعی، ساختواژه زبان فارسی به رده زبان‌های تحلیلی گرایش بیشتری دارد و به سوی تحلیلی‌تر شدن می‌رود.

**واژه‌های کلیدی:** صرف توزیعی، رده‌شناسی ساختواژی، رویکرد کمی، زبان تحلیلی، زبان پیوندی.

## ۱ - مقدمه

ساختواژه علمی است که به مطالعه چگونگی ساخت واژه و تناظر نظام‌مند صورت‌معنا می‌پردازد و ساختار درونی واحدهای زبانی را بررسی و تحلیل می‌کند. زبان‌ها به شیوه‌ای یکسان واژه نمی‌سازند و از ابزارهای صرفی نحوی مختلفی برای تولید واژه بهره می‌گیرند. در برخی از زبان‌ها اغلب واژه‌ها از یک تکواژ تشکیل می‌شوند و در گروهی دیگر تکواژهای سازنده واژه چنان درهم می‌آمیزند که نمی‌توان آنها را به سادگی از هم

تفکیک کرد. زبان‌های بسیاری نیز به لحاظ ساختواژی میان این دو طیف می‌نشینند. در این میان، رده‌شناسی ساختواژی به بررسی علمی و طبقه‌بندی زبان‌های بشری بر پایه مطالعات ساختواژی می‌پردازد و می‌کوشد تا از این منظر، شباهت‌ها و افتراق‌های گونه‌های زبانی را بیابد. با توجه به این‌که این نوع رده‌شناسی از یافته‌های ساختواژی مایه می‌گیرد، باید بر یکی از رویکردهای نظری علم ساختواژه مبتنی باشد. مقاله حاضر ضمن بازنگری در برخی از تحلیل‌های صرف سنتی، انگاره دستوری صرف توزیعی را به عنوان شالوده نظری خود برگزیده است تا از این رهگذر، به تحلیل و درونداد مناسبی برای رده‌شناسی ساختواژی زبان فارسی دست یابد.

صرف توزیعی، در تقابل با رویکردهای واژه‌گرا، بر آن است تا تحلیلی یکسان از ساخت سازه‌های مرکب (واژه، گروه و جمله) به دست دهد. هله و مرنتر (۱۹۹۳: ۱۱۲) تصریح می‌کنند که در این رویکرد، نحو به جای آن‌که با واژه‌های شکل‌یافته سروکار داشته باشد، با استفاده از مشخصه‌های صوری، سازه‌ها را تولید می‌کند. به گفته دقیق‌تر، این رهیافت نظری ادعا می‌کند که ساخت‌های تحلیلی، ترکیبی و بس‌ترکیبی (چندترکیبی) در همه زبان‌های گسسته، پیوندی و هم‌میخته در حوزه نحو تولید می‌شوند و بنابراین، سازوکارهای نحوی باید چنان کارآمد باشند که بتوانند از عهده تبیین این تنوعات زبانی برآیند. پرسش اصلی پژوهش حاضر این است که زبان فارسی بر پایه یافته‌های نظریه صرف توزیعی در کدام رده ساختواژی می‌گنجد و آیا انگاره دستوری صرف توزیعی نیز همچون رهیافت صرف واژگرا و سنتی، رده ساختواژی زبان فارسی را جزو زبان‌های تحلیلی قرار می‌دهد و مؤید نتایج پژوهش‌های پیشین در این زمینه است. به این ترتیب، پژوهش پیش رو می‌کوشد تا با تمرکز بر دو پرسش زیر، تحلیل تازه‌ای از تعیین رده ساختواژی زبان فارسی در چارچوب صرف توزیعی ارائه کند: پرسش اول: جایگاه زبان فارسی بر روی محور پیچیدگی درون ساخت‌های دستوری کجاست؟

پرسش دوم: از منظر شفافیت مرز میان تکواژها، زبان فارسی در چه رده زبانی می‌گنجد؟

نکته مهم دیگر این است که احتمالاً هیچ زبانی نیست که دارای رده زبانی ساختواژی محض باشد. در واقع، هیچ زبانی مطلقاً گسسته، پیوندی یا هم‌میخته نیست، بلکه رده ساختواژی زبان‌ها بر اساس گرایش غالبی که در ساختواژه آن زبان وجود دارد، تعیین می‌شود (سایپر، ۱۹۲۱: ۱۲۳). پیش‌بینی می‌شود زبان فارسی نیز این‌گونه باشد و به یکی از رده‌های زبانی گرایش بیشتری داشته باشد. بر این اساس، این مقاله در چهار

بخش تنظیم شده است. پس از مقدمه حاضر در بخش دوم، مبانی نظری پژوهش را در دو زیر بخش معرفی خواهیم کرد. در بخش سوم به پیشینه پژوهش می‌پردازیم و در بخش چهارم نیز به منظور بررسی رده ساختارهای زبان فارسی، داده‌های پژوهش را تحلیل و بررسی می‌کنیم. سرانجام، در فصل پنجم نشان می‌دهیم که مطابق با پرسش نخست پژوهش، جایگاه زبان فارسی بر روی محور پیچیدگی درون ساختارهای دستوری واژه‌ها، به رده زبان‌های تحلیلی نزدیک است و مطابق با پرسش دوم پژوهش، زبان فارسی از منظر شفافیت مرز میان تکواژها، به رده زبان‌های پیوندی تعلق دارد.

## ۲ - مبانی نظری پژوهش

بخش حاضر در دو زیر بخش مستقل مبانی نظری پژوهش را به اجمال معرفی می‌کند. زیر بخش نخست به رده‌شناسی ساختارهای زبان‌ها و معیارهای رده‌بندی زبان‌ها بر پایه فرایند تولید واژه و چگونگی شکل‌گیری ساختارهای دستوری می‌پردازد. زیر بخش دوم با اشاره‌ای به کلیات نظریه صرف توزیعی، تحلیلی از تکواژهای تصریفی و اشتقاقی به‌دست می‌دهد که متفاوت با رویکرد صرف سنتی خواهد بود.

### ۲-۱- رده‌شناسی ساختارهای

از انواع مطالعات رده‌شناسی زبان می‌توان به رده‌شناسی ساختارهای زبان اشاره کرد که خاستگاه آن به سده نوزدهم میلادی بازمی‌گردد. رده‌شناسی ساختارهای زبان رویکردی کل‌گرا به زبان دارد و نظام تصریفی زبان را بررسی و مطالعه می‌کند. در رویکرد یادشده بر اساس ویژگی‌های دستگاه صرفی زبان، طبقه‌بندی ساختارهای صورت می‌گیرد. در یافته‌های فریدریش فن اشلگل (۱۸۰۸) می‌توان نخستین دستاوردهای رده‌شناسی ساختارهای را جست. وی زبان‌ها را به دو گروه زبان‌های وندی و زبان‌های تصریفی تقسیم کرد. آگوست فن اشلگل (۱۸۱۸) دسته‌بندی دوگانه موجود را به دسته‌بندی سه‌گانه زبان‌های بدون ساخت دستوری، زبان‌های وندی و زبان‌های تصریفی تغییر داد. آگوست همچنین رده‌بندی جزئی‌تری از زبان‌های تصریفی با عنوان زبان‌های ترکیبی و زبان‌های تحلیلی ارائه کرد. سپس ویلهلم فن هومبالت (۱۸۲۵ و ۱۸۳۶) به سه رده اشلگل، زبان‌های انضمامی را افزود و زبان‌ها را به چهار دسته گسسته، تصریفی، پیوندی و انضمامی تقسیم کرد. ساپیر از رده‌های زبانی گسسته، پیوندی، تصریفی و بس‌ترکیبی نام می‌برد و عنوان می‌کند رده یک زبان را نمی‌توان با قطعیت تعیین کرد و در تعیین رده زبانی، گرایش کلی و غالب آن حائز اهمیت است (ساپیر، ۱۹۲۱: ۱۲۳). ساپیر ویژگی‌های ساختارهای را بر اساس دو معیار کلی ارائه کرد و نشان داد که زبان‌ها از جهات مختلف قابل رده‌بندی هستند. او ابراز داشت بررسی ساخت درونی واژه‌ها از دو جنبه قابل طرح است که هر یک به نوبه خود رده‌هایی از زبان‌ها را دسته‌بندی می‌کنند. این دو عامل عبارت‌اند از: شفافیت مرز ساختارهای تکواژها، و میزان پیچیدگی درونی

ساخت‌های دستوری واژه‌ها که بر تعداد تکواژهای موجود در هر واژه دلالت دارد (ساپیر، ۱۹۲۱؛ رضایی، بهرامی، ۱۳۹۴: ۷). ساپیر عامل نخست را که در واقع، فرایندهای صوری است و مبتنی بر شیوه ساخته شدن واژه‌ها در یک زبان است، شیوه نامیده و بر اساس آن، سه رده زبانی را بازشناخته است. شاخص دوم ساپیر که مکمل معیار نخست قلمداد می‌شود، استحکام نسبی پیوند میان عناصر وندی با هسته واژه است؛ یعنی در این ملاک، درجه ترکیب‌پذیری واژه در یک زبان مطرح است. و این ملاک را ترکیب خوانده و بر اساس آن، سه رده زبانی‌های تحلیلی، زبان‌های ترکیبی و زبان‌های بس‌ترکیبی را بازشناخته است (گرینبرگ، ۱۹۷۴: ۱۲۶-۱۴۲؛ دبیرمقدم، ۱۳۹۲: ۱۵-۱۷).

#### ۲-۱-۱- معیار شفافیت مرز تکواژها

بر اساس این معیار، ساپیر سه رده زبانی گسسته، پیوندی و هم‌آمیخته را بازشناخته است. الف - زبان‌های گسسته: در زبان‌های گسسته، تناظر یک به یک میان تکواژها و واژه‌ها وجود دارد. این سخن بدین معناست که در چنین زبان‌هایی معمولاً هر تکواژ برابر با یک واژه است. ویتنامی نمونه بسیار مناسبی از زبان‌های گسسته و فاقد ساختواژه است (ساپیر، ۱۹۲۱: ۶۱-۶۳؛ آیکنوالد، ۳/۲۰۰۷: ۳؛ تامپسون ۱۹۸۷: ۲۰۷).

ب - زبان‌های پیوندی: در زبان‌های پیوندی هر واژه می‌تواند از چند تکواژ تشکیل شود، با این‌همه مرز میان تکواژها روشن و مشخص است و به‌ندرت تکواژ آمیخته‌ای در این زبان‌ها یافت می‌شود. در این گروه از زبان‌ها، برخلاف زبان‌های گسسته، معمولاً تناظر یک به یک میان واژه و تکواژ وجود ندارد، گو این‌که ممکن است برخی از واژه‌ها یک تکواژی باشند. در زبان‌های پیوندی هر تکواژ صرفاً حامل یک معناست و به‌همین دلیل، این عناصر زبانی را می‌توان به‌سادگی از هم تفکیک کرد. برای مثال، اسم در ترکی و مجاری که از مشهورترین زبان‌های پیوندی جهان‌اند، به آسانی به ستاک واژگانی و مجموعه‌ای از وندها قابل تقسیم است (ساپیر، ۱۹۲۱: ۶۱-۶۴؛ آیکنوالد، ۳/۲۰۰۷: ۴).

پ - زبان‌های هم‌آمیخته: در زبان‌های هم‌آمیخته هر واژه می‌تواند از چند تکواژ تشکیل شود، اما برخلاف زبان‌های پیوندی، مرز میان تکواژها مشخص نیست و عناصر سازنده واژه را نمی‌توان به‌سادگی از هم تفکیک کرد. در زبان‌های هم‌آمیخته که روسی، صربی، کروات و لاتین نمونه‌های مشهور آن به‌شمار می‌روند، یک تکواژ یا گاهی آمیزه‌ای از چند تکواژ بیش از یک معنای دستوری یا واژگانی را بر خود حمل می‌کند (ساپیر، ۱۹۲۱: ۶۱-۶۴؛ آیکنوالد، ۳/۲۰۰۷: ۴).

#### ۲-۱-۲- معیار پیچیدگی درونی ساخت‌های دستوری واژه‌ها



چنان که پیشتر اشاره کردیم، بر پایه معیار دوم، یعنی میزان پیچیدگی درونی ساخت‌های دستوری واژه‌ها، سه گونه زبان قابل شناسایی‌اند: زبان‌های تحلیلی، زبان‌های ترکیبی و زبان‌های بس ترکیبی (چند ترکیبی).

الف - زبان‌های تحلیلی: در زبان‌های تحلیلی مفاهیم دستوری، نه با وندافزایی، بلکه از رهگذر واژه‌های مستقل بیان می‌شوند و در نتیجه، میزان پیچیدگی درونی ساخت‌های دستوری واژه‌ها بسیار اندک است. بر این اساس، به درستی انتظار می‌رود که همه زبان‌های گسسته، از جمله چینی و ویتنامی، در گروه زبان‌های تحلیلی قرار بگیرند، زیرا چنین زبان‌هایی عمدتاً فاقد وند و تکواژ وابسته‌اند و مفاهیم دستوری را با واژه‌های منفرد بیان می‌کنند (ساپیر، ۱۹۲۱: ۶۱-۶۲؛ آیکنوالد، ۲۰۰۷: ۳/۷).

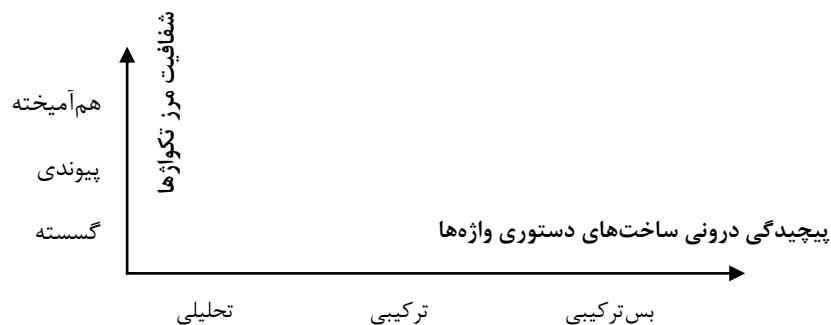
ب - زبان‌های ترکیبی: در زبان‌های ترکیبی هر واژه معمولاً متشکل از چند تکواژ است، زیرا در این زبان‌ها مفاهیم دستوری از طریق وندها و عناصر وابسته صرفی بیان می‌شوند. به همین دلیل، میزان پیچیدگی درونی ساخت‌های دستوری واژه‌ها نسبت به زبان‌های تحلیلی بیشتر است (ساپیر، ۱۹۲۱: ۶۱-۶۴؛ آیکنوالد، ۲۰۰۷: ۳/۷).

پ - زبان‌های بس ترکیبی: در زبان‌های بس ترکیبی یا چند ترکیبی، واژه‌ها بیشترین پیچیدگی را در درون ساختار خود دارند و مفاهیم متعددی که در زبان‌های تحلیلی و حتی ترکیبی به صورت گروه یا جمله بیان می‌شوند، در قالب واژه درمی‌آیند. این ویژگی نتیجه فرایند انضمام عناصر مختلف واژگانی و دستوری و تبدیل آنها به یک واژه واحد است. به همین دلیل، چنین زبان‌هایی را انضمامی نیز می‌خوانند. زبان‌های بس ترکیبی ممکن است پیوندی باشند و عناصر منضم شده مرز مشخصی داشته باشند، یا می‌توانند هم آمیخته باشند (ساپیر، ۱۹۲۱: ۶۱-۶۴؛ آیکنوالد، ۲۰۰۷: ۳/۸). در مجموع، زبان‌های بس ترکیبی معمولاً دارای تکواژهای وابسته فراوان و شمار اندکی ستاک‌های مستقل‌اند. فرایند انضمام در این زبان‌ها پربسامد است و اسم‌ها و قیدها در درون گروه‌های فعلی، و صفت‌ها در درون گروه‌های اسمی منضم می‌شوند و بدین ترتیب، واژه‌هایی با ساختار درونی پیچیده تولید می‌کنند (آیکنوالد، ۲۰۰۷: ۳/۸).

تمایز میان زبان‌های تحلیلی، ترکیبی و بس ترکیبی بر روی یک پیوستار قرار دارد؛ بدین معنا که مرز مشخص و معینی میان رده‌های زبانی یادشده وجود ندارد. آیکنوالد (۲۰۰۷: ۳/۷) به پیروی از گرینبرگ (۱۹۵۴) که پژوهش‌های مختلف، روش آن را پذیرفته‌اند، یادآور می‌شود که با تقسیم تعداد تکواژهای هر جمله بر تعداد واژه‌ها، می‌توان میزان تحلیلی یا ترکیبی بودن هر زبان را محاسبه کرد. افزون بر این، آیکنوالد (۲۰۰۷: ۳/۸) تصریح می‌کند که دو معیار پیش‌گفته شاخص‌های رده‌شناختی مستقلی‌اند که برای هر زبان باید جداگانه بررسی شوند. مثلاً، زبان‌های گسسته تحلیلی‌اند، اما برعکس آن الزاماً صحیح نیست. دو معیار ترکیب‌پذیری و شفافیت مرز



تکواژها گرچه با یکدیگر در تعامل اند، معیارهای مستقلی هستند. آیکنوالد نمودار دوبعدی زیر را برای به‌کارگیری هم‌زمان، اما مستقل این دو معیار و تعامل بین آنها پیشنهاد می‌دهد (۲۰۰۷: ۳: ۷).



گرینبرگ (۱۹۵۴) با پذیرش مطالعات ساپیر، یک گام پیش گذاشت و شاخص کمی در تعیین میزان ترکیب‌پذیری زبان‌ها ارائه داد. این شاخص از طریق تقسیم تعداد تکواژهای یک حجم نمونه، به تعداد واژه‌های موجود در آن به‌دست می‌آید. بر اساس معیار کمی که گرینبرگ بیان می‌کند، زبان‌ها به سه دستهٔ تحلیلی و ترکیبی و بس ترکیبی تقسیم می‌شوند (گرینبرگ، ۱۹۵۴: ۱۹۴):

(۱). الف) اگر زبانی بین ۱ تا ۱,۹۹ صدم تکواژ در هر واژه داشته باشد، تحلیلی است.

(ب) اگر زبانی بین ۲ تا ۲,۹۹ صدم تکواژ در هر واژه داشته باشد، ترکیبی است.

(پ) اگر زبانی بیشتر از ۳ تکواژ در هر واژه داشته باشد، بس ترکیبی یا چندترکیبی است.

بدیهی است هرچه حاصل تقسیم به حداکثر محدوده نزدیک‌تر باشد، تمایل زبان برای داشتن ویژگی‌های رده بعد از خود افزایش می‌یابد. در ادامه، بررسی‌ها به منظور تعیین ردهٔ ساختواژی زبان فارسی از دیدگاه صرف توزیعی پی گرفته می‌شود.

## ۲-۲- صرف توزیعی

در انگارهٔ صرف توزیعی همهٔ عناصر مرکب در نحو ساخته می‌شوند و تکواژها و مشخصه‌های صوری درون‌داد نهند و پایانه‌های نحوی را تشکیل می‌دهند. چنین رهیافتی به زبان در تقابل با کمینه‌گرایی واژه‌گرا قرار دارد که طبق آن، نظام زایشی غیرنحوی مسئول اشتقاق واژه‌هاست. چهار ویژگی بنیادین در نظریهٔ صرف توزیعی وجود دارد که سبب تمایز آن با کمینه‌گرایی واژه‌گرا می‌شود (صدیقی ۲۰۰۹؛ امبیک ۲۰۱۵):  
الف - مقوله‌سازی: ریشه‌های واژگانی بدون مقولهٔ دستوری وارد اشتقاق می‌شوند. نحو مقولهٔ دستوری ریشه‌ها را تعیین می‌کند.

ب - درج مؤخر: عناصری که دستخوش فرایندهای نحوی می‌شوند، انتزاعی و فاقد جوهر آوایی‌اند. درج جوهر آوایی در گره‌های پایانی، پس از نحو و در سطح آوایی میسر می‌شود.

پ - تجزیه صرفی نحوی: هیچ صورت غیربسیطی در حافظه ذخیره نمی‌شود؛ همه سازه‌های مرکب هر بار که به کار می‌روند، توسط دستور ساخته می‌شوند.

ت - زیرتخصیصی: عناصر واژگانی و دستوری با کمترین مشخصه‌های صرفی نحوی ممکن ذخیره می‌شوند. مشخصه‌های این عناصر باید زیرمجموعه‌ای از گره هدف باشد. در این نظریه، مفهوم ریشه با تعبیر سنتی این اصطلاح متفاوت است. در انگاره صرف توزیعی، ریشه مفهومی انتزاعی است، و نه واژه یا تکواژی عینی. برای مثال، «ساختن» مفهومی انتزاعی در ذهن اهل زبان است که در واژه‌های متعددی مثل «می‌سازد»، «ساخت‌وساز»، «ساختار» و «ساختمان» تبلور می‌یابد. این معنای مشترک همان ریشه یا مفهوم انتزاعی «ساختن» است که طبق قرارداد، با نماد «ساز/√» نمایش داده می‌شود. در اینجا مفهوم انتزاعی ریشه با تعبیر سنتی آن متفاوت است.

گره‌های نحوی در اشتقاق از مشخصه‌های صوری و ریشه‌ها تشکیل می‌شوند و هنگامی که همه فرایندهای نحوی و از جمله ادغام و حرکت پایان می‌پذیرند، محتوای واجی مناسب به گره‌های پایانی اشتقاق افزوده می‌شود. این محتوای واجی در بخشی از دستور به نام واژگانه ذخیره می‌شود (به نقل از انوشه ۱۳۹۷). واژگانه فهرستی از عناصری است که نقش‌شان در دستور فراهم آوردن جوهره آوایی برای ریشه‌ها و مشخصه‌هایی است که در گره‌های پایانی اشتقاق نحوی حضور دارند. از این طریق، یعنی با درج عناصر متعلق به واژگانه در اشتقاق، ساخت نحوی می‌تواند به تلفظ درآید. هر عنصری که در درون این فهرست قرار دارد، واحد واژگانه‌ای، یا به اختصار واژ، خوانده می‌شود. هر واژ ممکن است به اشکال آوایی مختلف تظاهر یابد که به آنها واژگونه می‌گویند. مثلاً، «ساز/√» دارای دو واژگونه /sāz/ و /sāx/ است:

(۲). مشخصه‌های «ساز/√»:

الف. /sāx/	[√Sāz, +t]	ساخت، ساختار، ساختن، ساخته، ساختمان
ب. /sāz/	[√Sāz]	می‌سازد، سازگار، سازش، سازه، سازمان

واژگونه (الف) در گرهی درج می‌شود که اولاً ریشه «ساز/√» در آن باشد و ثانیاً وندی در مجاورت آن حضور داشته باشد که با همخوان /t/ آغاز شود. واژگونه (ب) نیز تابع شرط دگر جای است. یعنی، در هر هدفی که ریشه «ساز/√» حضور داشته باشد و واژگونه (الف) مجاز به درج نباشد، واژگونه (ب) حضور می‌یابد.

صرف توزیعی از سه فرایند پسانحوی برای در هم آمیختن یا تفکیک گره‌ها استفاده می‌کند که عبارت‌اند از: ادغام صرفی، همجوشی و شکافت. طبق مفهوم ادغام صرفی، در هر سطحی از تحلیل نحوی (ژساخت، رساخت، سطح واجی) رابطه میان X و Y

می‌تواند با وندافزایی هسته واژگانی X به هسته واژگانی Y جایگزین (بیان) شود. فرایند همجوشی واکنشی هسته‌ای است که در پی آن دو هسته خواهری که پیشتر با اعمال ادغام صرفی ذیل یک گره قرار گرفته‌اند، در قالب یک گره پایانی واحد به هم می‌پیوندند. در مقابل همجوشی که دو گره پایانی مجاور را درمی‌آمیزد و به یک گره واحد تقلیل می‌دهد، فرایند شکافت قرار دارد که یک گره را به دو یا چند گره تقسیم می‌کند. نخستین بار نویر (۱۹۹۷) و مرنتز (۱۹۹۷) این فرایند را معرفی کرده و آن را در گسستن گره مطابقه به چند گره مستقل به کار بسته‌اند. در پی اعمال فرایند شکافت بر هسته مطابقه، مشخصه‌های شخص و شمار (و گاهی جنس دستوری) دارای گره‌های مستقلی می‌شوند و هریک با واژ جداگانه‌ای تحقق می‌یابند. در بخش بعدی پیشینه پژوهش حاضر را معرفی می‌کنیم.

### ۳ - پیشینه پژوهش

پیشینه حاضر بر دو حوزه متمرکز شده است. ابتدا به آن دسته از مطالعات رده‌شناختی که با مقاله حاضر مرتبطند، اشاره می‌کنیم و سپس به پیروی از انوشه (۱۳۹۷) نشان می‌دهیم که انگاره صرف توزیعی چه تأثیری بر تحلیل تکواژی ساخت‌های زبان فارسی به جای می‌گذارد.

#### ۳ - ۱ - مطالعات رده‌شناسی ساختواژی

رده‌شناسی زبان که مطالعه نظام‌مند تنوع بین زبان‌ها تعریف می‌شود (کامری، ۲۰۰۱: ۲۵)، قدمتی ۲۵۰ ساله دارد و از پیش از سده بیستم تاکنون، زبان‌شناسانی که رویکرد کل‌گرا به آن داشته‌اند؛ همچون گابریل جرارد، فریدریش فن اشلگل، آگوست فن اشلگل، ویلهلم فن هومبلت، آگوست اشلايخر، ادوارد ساپیر و ولادیمیر اسکالیچیکا هریک از جنبه‌های مختلف رده‌بندی کلی از زبان‌ها به دست داده‌اند و پژوهش‌های بسیاری بر اساس آنها به ثمر رسیده است. از آن میان، آیکنوالد (۳/۲۰۰۷: ۳) در پژوهش خود با عنوان «تمایزهای رده‌شناختی در واژه‌سازی» بیان داشته است که بر اساس دو معیار شفافیت مرزهای صرفی میان تکواژهای سازنده واژه‌ها (محور عمودی) و نیز میزان پیچیدگی درونی ساخت‌های دستوری واژه‌ها (محور افقی) که ساپیر (۱۹۲۱) آن را برای تعیین رده یک زبان ارائه کرده است، رده ساختواژی زبان‌های مختلف را می‌توان تعیین کرد. وی در پژوهش خود با بررسی زبان‌های مختلف اذعان کرده است که زبان ویتنامی و زبان چینی باستان نمونه‌ای از زبان‌هایی هستند که بر اساس دو معیار یادشده، در محور دوبعدی بر روی مختصات [تحلیلی و گسسته] قرار می‌گیرند و زبانی با رده تحلیلی گسسته به شمار می‌روند. آیکنوالد در پژوهش خود با بررسی زبان مجارستانی،

عنوان کرده است که این زبان بر اساس دو معیار یادشده ساپیر، در محور دوبعدی بر روی مختصات [ترکیبی و پیوندی] قرار می‌گیرد که در واقع می‌توان گفت که زبان مجارستانی به رده زبان‌های ترکیبی پیوندی تعلق دارد. وی با بررسی زبان روسی، بیان کرده است که این زبان جزو زبان‌هایی تعیین شده که بر روی مختصات [ترکیبی و هم‌آمیخته] واقع می‌شود و زبان روسی را می‌توان جزو رده زبان‌های ترکیبی آمیخته دسته‌بندی کرد. بررسی زبان اسکیمویی در پژوهش آیکنوالد نشان می‌دهد که این زبان بر اساس دو معیار یادشده ساپیر بر روی مختصات [بس ترکیبی و پیوندی] قرار می‌گیرد. از این‌رو، زبان اسکیمویی جزو زبان‌های بس ترکیبی پیوندی رده‌بندی می‌شود. همچنین زبان سرخ‌پوستان آپاچی نیز در این پژوهش به رده زبان‌های بس ترکیبی هم‌آمیخته تعلق دارد (آیکنوالد، ۲۰۰۷: ۳/۸). از سوی دیگر، شماری از زبان‌شناسان و پژوهشگران زبان فارسی نیز به تعیین رده ساختواژی زبان فارسی در دوره‌های مختلف پرداخته‌اند. در این بین می‌توان به ثمره (۱۳۶۹) اشاره کرد که پس از آوردن نمونه‌هایی از شاخه‌های زبانی، دو متغیر عمده رده‌شناسی زبان، ساختواژه و آرایش جمله را معرفی می‌کند. وی به جای استفاده از یک متغیر کلی با عنوان ساختواژه، دو متغیر تعداد تکواژها در واژه و نیز حد تقطیع‌پذیری آنها را ملاک قرار می‌دهد. بر این اساس، دو پیوستار به دست می‌دهد؛ یکی پیوستار شمار تکواژها در واژه‌ها و دیگری پیوستار نحوه ترکیب تکواژها در داخل واژه. دو قطب پیوستار اول را زبان منفرد و زبان بساوندی (بس ترکیبی) و دو قطب پیوستار دوم را زبان پیوندی و زبان ترکیبی تشکیل می‌دهند. ثمره تعیین رده ساختواژی فارسی را بر اساس داده‌های خاصی انجام نداده است. او ویژگی‌های رده‌شناختی زبان فارسی را چنین برشمرده است؛ از لحاظ نسب: از خانواده زبان‌های ایرانی، از شاخه هندوایرانی، متعلق به تیره هندواروپایی. از لحاظ ساختواژه: زبان پیوندی با رنگ ضعیفی از ترکیبی. از لحاظ آرایش جمله: عمل - عامل.

در پژوهش شریفی و اخلاقی (۱۳۹۲) رده ساختواژی زبان فارسی بر اساس رویکرد کمی رده‌شناسی بررسی شده است. نتیجه میانگین مقدار کمی که از شمارش تکواژها در رویکرد کمی گرینبرگ حاصل شده، عدد ۱,۵۳ است که زبان فارسی را به لحاظ ساختواژی در رده زبان‌های تحلیلی قرار می‌دهد. گفتنی است که در این پژوهش رده ساختواژی زبان فارسی تنها در محور افقی که مربوط به «ترکیب‌پذیری واژه‌ها» است، بررسی شده و محور عمودی که در بردارنده «شفافیت مرز تکواژها» است، پرداخته نشده است. همچنین این‌که چند نکته قابل توجه در پژوهش یادشده دیده می‌شود. نخست این‌که به واژه‌هایی که در دستگاه ساختواژی غیر از فارسی ساخته شده‌اند، اشاره‌ای نشده و مشخص نشده است که وام‌واژه‌هایی همچون «نقاش»، «اصلح»، «مراکز» و «تجار» که دستگاه ساختواژی زبان عربی آنها را تولید کرده است، کنار گذاشته شده‌اند

یا همچون واژه‌های تولیدشده در دستگاه ساختواژه فارسی، این واژه‌ها تکواژشماری شده‌اند و واژ آمیخته به‌شمار رفته‌اند. دوم این‌که، در پژوهش شریفی و اخلاقی ذکر نشده است که تقطیع و شمارش تکواژهای واژه‌های فارسی بر اساس کدام رویکرد و انگاره صورت گرفته است. به نظر می‌رسد که این تقطیع تکواژی بر اساس صرف سنتی انجام شده باشد. انگاره واژه‌گرای صرف سنتی در تقطیع تکواژی، تفاوت‌های قابل توجهی با تقطیع تکواژی در انگاره ناواژه‌گرای صرف توزیعی دارد. برای مثال، واژه «خورده» در صرف سنتی از اتصال سه تکواژ «خور + د + ه» ساخته‌شده ولی از منظر صرف توزیعی از اتصال دو تکواژ «خور + ده» ساخته است. دبیرمقدم (۱۳۹۲: ۱۷) بیان کرده است که در طول تاریخ یک زبان، گاه دیده می‌شود که گرایشی به سوی تحلیلی‌تر شدن وجود دارد. وی اشاره کرده که زبان فارسی از دوره میانه به بعد دارای چنین گرایشی است و به سوی تحلیلی‌تر شدن پیش می‌رود. دبیرمقدم (۱۳۹۲: ۱۹) ویژگی‌های رده‌شناختی ساختواژی زبان فارسی امروز را تلفیقی از موارد زیر برمی‌شمرده است؛ داشتن جلوه‌هایی از زبان‌های تصریفی: مانند ستاک گذشته فعل‌های بی‌قاعده همچون «دوخت» در مقایسه با ستاک حال آن «دوز». برخورداری از جنبه‌هایی از زبان‌های پیوندی: مانند واژه «ناسازگاری». جایز شمردن انضمام: نظیر صورت غیرمنضم «دانش‌آموزان درس‌هایشان را خوب خواندند» و گونه منضم آن «دانش‌آموزان خوب درس خواندند». وجود ویژگی‌های مختلف زبان‌های تحلیلی: مانند وجود فعل‌های وجهی مستقل، بهره‌گیری از پیش‌افزافه. وی در نهایت، گرایش غالب زبان فارسی امروز را به سوی رده‌های تحلیلی دانسته است.

### ۳-۲- تقطیع تکواژی در صرف توزیعی

برای تعیین رده ساختواژی و تحلیل کمی زبان فارسی نیازمند شناسایی و تقطیع تکواژها هستیم. در میان تکواژهای فارسی، وندهای اشتقاقی بسیار شفاف‌تر از عناصر تصریفی‌اند و فرایند همجوشی در آنها کمتر رخ می‌دهد. در تحلیل انوشه (۱۳۹۷) بین ریشه‌های فعلی و وندهای تصریفی زمان گذشته، مصدرساز، و اسم مفعول فرایند همجوشی رخ نمی‌دهد و عناصر یادشده در فارسی به ترتیب «د/ت»، «دن/تن» و «ده/ته» هستند. این سخن بدین معناست که اولاً در فارسی فقط یک جفت وند ماضی‌ساز وجود دارد. ثانیاً وند مصدری متشکل از «د/ت + ن» نیست و «دن/تن» یک تکواژ واحد است. ثالثاً وند اسم مفعول یا صفت مفعولی نیز از دو جزء «د/ت + ه» ساخته نشده است. طبق این رهیافت نظری که با رویکرد تاریخی به دستور زبان نیز سازگار است، تقطیع سه صورت تصریفی ریشه «بین<sup>۱</sup>» به صورت زیر است:

(۳). صورت‌های تصریفی «بین<sup>۱</sup>»:

الف. دیدم	←	دید + د + م
ب. دیدن	←	دید + دن
پ. دیده	←	دید + ده

به این ترتیب، انوشه (۱۳۹۷) بر اساس رویکرد صرف توزیعی چنین ادعا کرده است که تنها وند گذشته در فارسی تناوب «د/ت» است و اگر چهار وند «دید»، «د/ت»، «ست» و «اد»، را واژگونه‌های ماضی‌ساز زبان فارسی بدانیم، از تبیین برخی از داده‌ها بازمی‌مانیم و حتی ناگزیر می‌شویم فهرست وندهای گذشته‌مان را گسترش دهیم. مثلاً، «فت» را هم باید وند گذشته‌ای تلقی کنیم که با اتصال به صورتی مانند «پذیر»، ستاک گذشته «پذیرفت» را می‌سازد. وی در تحلیل خود خاطر نشان کرده است که تقلیل وندهای گذشته فارسی به یک وند «د/ت» پیامدهای متعددی دارد: یک: بر اساس این رویکرد وندهای مصدرساز فارسی به جفت «دن/تن» و وندهای صفت مفعولی فارسی به جفت «ده/ته» محدود می‌شوند که هر دو جفت نیز پیش‌بینی‌پذیرند.

دو: در تحلیل‌های رایج از وند گذشته زبان فارسی، از میان چهار واژگونه «دید»، «د/ت»، «ست» و «اد»، غالباً وند گذشته «دید» صورت پایه‌ای تلقی می‌شود و مصادر جعلی و فعل‌های ساده‌ای که بر پایه اسم‌های فارسی یا قرضی ساخته می‌شوند (مانند: طلبیدن، شوتیدن، جهنمید)، در صورت‌های صرفی خود وند یادشده را می‌پذیرند. در تحلیل حاضر تنها وند ماضی‌ساز زبان فارسی تناوب «د/ت» است و اساساً وندی با صورت آوایی «دید» در این زبان وجود ندارد. به جای آن، اهل زبان به شکلی قاعده‌مند با افزودن واج پایانی /i/ به ریشه، واژگونه‌ای را می‌سازند که در زمان گذشته یا در ساختار مصدر به کار می‌رود. از آنجا که این واج پایانی از جنس واکه و طبیعتاً واکدار است، دو وند «د» و «دن»، به ترتیب، برای صورت گذشته و صورت مصدری جذب می‌شوند. سه: نکته مهم دیگر این که واژگونه‌هایی که برای ساخت صورت حال از یک سو و ستاک گذشته و مصدر از سوی دیگر به کار می‌روند، صرفاً محدود به این بافت‌های صرفی نیستند و در ساختار واژه‌های مرکب و مشتقی مانند صفت فاعلی و مفعولی، اسم مصدر و جز آنها نیز حضور می‌یابند. برای مثال، ریشه‌های «ل‌رو» و «ل‌ساز» واجد دو واژگونه مختلف‌اند که هر کدام در ساخت‌های مختلفی آشکار می‌شوند:

(۴). ریشه ل‌گو:

الف. واژگونه «گو»:	بگو،	می‌گویم،	گویش،	گویا،	گوینده،	واگوبه
ب. واژگونه «گف»:	گفتن،	می‌گفتم،	گفتار،	گفته،	گفتنی،	گفتمان

(۵). ریشه ل‌ساز:

الف. واژگونه «ساز»:	بساز،	می‌سازد،	سازش،	سازگار،	سازمان،	سازه
ب. واژگونه «ساخت»:	ساختن،	می‌ساخت،	ساختار،	ساخته،	ساختمان،	ساخت‌وساز

چهار: رویکردی که وندهای گذشته زبان فارسی را از چهار به یک تقلیل می‌دهد، بهتر می‌تواند نوع وند گذشته را در فعل‌های موسوم به سببی توضیح دهد. در داده‌های بالا و همچنین در ریشه‌ای مثل «شناس<sup>۱</sup>» با دو واژگونه مواجه‌ایم: «شناخ» و «شناس». در همه این نمونه‌ها، یک واژگونه در کنار وندی می‌نشیند که با همخوان دندانی /D/ آغاز می‌شود (یعنی /d/ یا /t/). واژگونه دیگر که بی‌نشان است، طبق شرط دگرجایی، در دیگر بافت‌ها درج می‌شود. برای مبنای گفت که در «شناخت» به دلیل مجاورت «شناس<sup>۱</sup>» با وند «ت» ماضی‌ساز، واژگونه «شناخ» پدیدار شده است. در ساخت گذشته سببی، یعنی «شناساند»، حضور وند سببی، سبب شده که دیگر ریشه در مجاورت وند دارای همخوان دندانی نیاید. به همین دلیل، به جای واژگونه «شناخ»، گونه «شناس» امکان درج مؤخر یافته است.

#### ۴ - تحلیل داده‌های ساختواژی

در بخش پیش رو به منظور بررسی رده ساختواژی زبان فارسی، نخست، از دیدگاه صرف توزیعی و به پیروی از رویکرد انوشه (۱۳۹۷) به تقطیع تکواژی واژه‌های پیکره زبانی که در دستگاه ساختواژی زبان فارسی ساخته شده‌اند، می‌پردازیم. سپس در بخش بعدی با استناد به یافته‌های انگاره دستوری صرف توزیعی، رده ساختواژی زبان فارسی را بر پایه معیارهای سایپر و نیز رویکرد کمی گرینبرگ بررسی و تحلیل خواهیم کرد.

#### ۴-۱ - تقطیع تکواژی واژه‌های پیکره

واژه از بنیادی‌ترین مفاهیم زبانی است و اهل زبان مرز آن را از یک‌سو با تکواژ و از سوی دیگر با گروه نحوی مشخص می‌کنند و خود از اجزای سازنده کوچک‌تری تشکیل می‌شود که این اجزای سازنده، تکواژ نامیده می‌شوند. بلومفیلد (۱۹۳۳: ۱۶۱) تکواژ را صورتی زبانی دانسته که هیچ شباهت آوایی-معنایی به دیگر صورت‌ها نداشته باشد. نزد هاکت (۱۹۵۸: ۱۲۳)، تکواژها کوچک‌ترین عناصر معنادار مستقل‌اند که در گفتار به کار می‌روند. تکواژ واحدی از صورت و معناست که یا محتوای معنایی دارد یا دستوری. تکواژها به دو گروه آزاد و وابسته تقسیم می‌شوند. تکواژهای آزاد می‌توانند مستقلاً در نقش واژه به کار روند یا همراه با دیگر تکواژها در ساخت واژه‌های تازه مشارکت کنند. در مقابل، تکواژهای وابسته هویت مجزا و مستقلاً ندارند و حضورشان به دیگر تکواژها وابسته است. گفتنی است تکواژهای آزاد و وابسته، تمایز میان تصریف و اشتقاق را رقم می‌زنند و تکواژهای تصریفی همواره مقولات دستوری را رمزگذاری می‌کنند؛ مانند شخص‌شمار، زمان و حالت دستوری. به منظور شناسایی تکواژهای اشتقاقی و تصریفی زبان فارسی و سهولت در تجزیه واژه‌ها به عناصر سازنده‌شان، از



اصول نایدا (۱۹۴۹ و ۱۹۶۵) بهره می‌گیریم. در ادامه، برخی از واژه‌های موجود در پیکره مورد بررسی را از دیدگاه ناواژه‌گرای صرف توزیعی تقطیع تکواژی می‌کنیم و جلوه‌هایی از ساخت‌های صرفی نحوی زبان فارسی را ارائه و دسته‌بندی می‌کنیم (انوشه، در دست چاپ): در زبان فارسی «ساخت آینده» به صورت تحلیلی و با فعل کمکی «خواستن» ساخته می‌شود؛ یعنی مفهوم دستوری، نه با وندافزایی، بلکه از رهگذر واژه‌ای مستقل بیان می‌شود. همچنین در زبان فارسی ساخت‌های ماضی بعید و التزامی نیز به صورت تحلیلی و با به‌کارگیری فعل کمکی ساخته می‌شوند. برای نمونه، فعل «خواهم فروخت»، دو واژه به‌شمار می‌رود که از چهار تکواژ ساخته می‌شود یا فعل «خورده بودیم» دو واژه تلقی می‌شود که در بردارنده پنج تکواژ است:

(۶). الف. «خواهم فروخت»: از اتصال تکواژهای (خواه + م + فروش + تن) ساخته می‌شود.

ب. «خورده بودیم»: از اتصال تکواژهای (خور + ده + باش + د + یم) ساخته می‌شود.

صورت فعلی گذشته در زبان فارسی به صورت ترکیبی ساخته می‌شود. گذشته ساده در زبان فارسی از طریق افزودن وند «د/ت» به ریشه فعلی شکل می‌گیرد؛ مانند ساخت گذشته ساده در واژه‌های «خوردم» و «برنمی‌گرداندم». واژه «خوردم»، از ریشه «خور» و وند گذشته‌ساز «-د» ساخته شده است. بنابراین، هریک از واژه‌های یادشده یک واژه به‌شمار می‌روند که در درون خود متشکل از چند تکواژ هستند و مفاهیم دستوری در آنها از طریق وندها و عناصر وابسته صرفی بیان شده‌اند. برای نمونه:

(۷). الف. «برنمی‌گرداندم»: از اتصال تکواژهای (بر + ن + می + گرد + ان + د + م) ساخته می‌شود.

ب. «خوردم»: از اتصال تکواژهای (خور + د + م) ساخته می‌شود.

ساخت‌هایی همچون «همیشه از خدا خواسته»، «خود بیمار انگاری»، «آماده به خدمت»، «جان در کف دست گرفته»، «از دماغ فیل افتاده»، «از خود راضی» و «پای در هوا» نمونه‌ای از زبان‌های بس‌ترکیبی یا چندترکیبی در زبان فارسی هستند که پیچیدگی زیادی در درون ساختار خود دارند. ساخت‌های یادشده در نتیجه فرایند انضمام عناصر مختلف واژگانی و دستوری و تبدیل آنها به یک واژه واحد ساخته شده‌اند. در این ساخت‌ها مفاهیم متعددی که در زبان‌های تحلیلی و حتی ترکیبی به صورت گروه یا جمله بیان می‌شوند، در قالب واژه مرکب آمده‌اند و یک واژه تلقی می‌شوند. برای نمونه، صفت‌های «آماده‌به‌خدمت» یا «ازخداخواسته» که از طریق فرایند انضمام تولید شده‌اند، یک واژه به‌شمار می‌روند که دارای تکواژهای زیر هستند:

(۸). الف. «آماده‌به‌خدمت»: از اتصال تکواژهای (آماده + به + خدمت) ساخته می‌شود.

ب. «ازخداخواسته»: از اتصال تکواژهای (از + خدا + خواه + ته) ساخته می‌شود.

ساخت‌هایی همچون «درست در کنار دیوار اتاق» و «با چند کتاب در دست» نمونه‌ای از زبان‌های گسسته در زبان فارسی هستند. در این نوع ساخت‌ها تناظر یک به یک میان

تکواژها و واژه‌ها وجود دارد و هر تکواژ برابر با یک واژه است. برای نمونه «درست در کنار دیوار اتاق» پنج واژه به‌شمار می‌رود که از پنج تکواژ ساخته شده است.

(۹). الف. «درست در کنار دیوار اتاق»: از اتصال تکواژهای (درست + در + کنار + دیوار + اتاق) ساخته می‌شود. ب. «با چند کتاب در دست»: از اتصال تکواژهای (با + چند + کتاب + در + دست) ساخته می‌شود.

واژه‌هایی همچون «دانشمندهایشان»، «برنمی‌گردانند» و «پژوهشگران» نمونه‌ای از ردهٔ زبان‌های پیوندی در زبان فارسی هستند که هریک از واژه‌ها از چند تکواژ تشکیل شده‌اند، با این‌همه مرز میان تکواژها روشن و مشخص است. برای مثال، واژه «دانشمندهایشان» یک واژه به‌شمار می‌رود که از تکواژهای «دان + ش + مند + ها + یشان» ساخته شده است. به‌نظر می‌رسد که ساختوارهٔ زبان فارسی نیز عمدتاً از این الگو پیروی می‌کند. در فارسی اغلب وندهای تصریفی و اشتقاقی تنها حامل یک معنای دستوری یا واژگانی هستند و با حفظ مرزهای خود به ریشه متصل می‌شوند. نمونه‌های زیر هریک یک واژه به‌شمار می‌روند که از پیوند وندها ساخته شده‌اند:

(۱۰). الف. «دبیرستانی‌ها»: از اتصال تکواژهای (دبیر + ستان + ی + ها) ساخته می‌شود. ب. «پراکندگی»: از اتصال تکواژهای (پراکن + ده + گی) ساخته می‌شود.

در ساختوارهٔ زبان فارسی هم‌آمیختگی وندها و ریشه‌ها بسامد اندکی دارد؛ برای نمونه، فعل سوم شخص مفرد «است» حاصل از آمیختگی ریشهٔ «باش» و زمان حال و وند صفر شناسه است. نمونه‌هایی از آمیختگی در زبان فارسی در ادامه آمده است:

(۱۱). الف. «رفته بودم»: از اتصال تکواژهای (رو + ته + باش + د + م) ساخته می‌شود. ب. «رفته است»: از اتصال تکواژهای (رو + ته + باش + 0) ساخته می‌شود. پ. «رفته‌ام»: از اتصال تکواژهای (رو + ته + باش + ام) ساخته می‌شود. ت. «رفته باشند»: از اتصال تکواژهای (رو + ته + باش + ند) ساخته می‌شود.

همان‌گونه که گفته شد، بسامد آمیختگی تکواژها در زبان فارسی بسیار پایین است و اغلب واژه‌های فارسی در دستگاه صرفی این زبان به صورت پیوندی تولید می‌شوند و وجود جلوه‌هایی از آمیختگی در زبان فارسی تأثیر چندانی در تغییر ردهٔ زبان فارسی به سوی ردهٔ زبان‌های آمیخته ایجاد نمی‌کند. وانگهی، بسیاری از واژه‌های آمیختهٔ فارسی نیز قرضی‌اند و بر مبنای قواعد ساختواژی زبان فارسی تولید نشده‌اند؛ مانند جمع‌های مکسری که از عربی به فارسی راه یافته‌اند. واژهٔ «کُتُب» در فارسی و عربی که افزون بر محتوای واژگانی، دارای مشخصهٔ دستوری جمع نیز هستند، به‌نوعی آمیخته (واژ آمیخته) محسوب می‌شوند. اما نکتهٔ مهم در پژوهش حاضر این است که واژه‌هایی که در دستگاه ساختواژی زبان فارسی دستخوش تصریف نشده‌اند، به‌هیچ‌وجه تکواژشماری و تحلیل نمی‌شوند؛ زیرا در رده‌شناسی ساختواژی، ویژگی‌های نظام صرفی

زبان مورد نظر تحلیل و تبیین می‌شود و وام‌واژه‌هایی که محصول ساختارهای زبان دیگری‌اند، از داده‌ها کنار گذاشته می‌شوند. بدین ترتیب، در میان داده‌های مقاله حاضر، نه واژه عربی «تجار» یافت می‌شود، نه حتی واژه «کفاش» (بر وزن فعال) که از ریشه فارسی «کفش»، بر پایه الگوهای ساختارهای زبان عربی تولید شده، لحاظ شده است. نکته دیگر این که بر پایه تحلیل نظریه صرف توزیعی، «دیدم» (و نمونه‌های مشابه، مانند «آدمم/ زدم») از سه تکواژ مستقل «دید + م + د + م» تشکیل شده است. بدین ترتیب، برخلاف تحلیل‌های سنتی (از جمله شریفی و اخلاقی ۱۳۹۲) در هیچ‌یک از فعل‌های زبان فارسی، ریشه با وند گذشته آمیخته نمی‌شود و واژ آمیخته شکل نمی‌گیرد. وانگهی، همان‌گونه که در داده‌های (۳) نشان دادیم، وندهای مصدری و صفت مفعولی در فارسی، به ترتیب، «دن/ تن» و «ده/ ته» هستند و این تکواژها را نباید دوجزئی به‌شمار آورد. برآیند نکات پیش‌گفته تأثیر مهمی در تقطیع تکواژی داده‌ها می‌گذارد. برای مثال، تحلیل تکواژی جمله ساده زیر را از دو منظر سنتی و صرف توزیعی مقایسه کنید (برگرفته از انوشه، در دست چاپ):

(۱۲). شعرا آمده بودند.

الف. صرف سنتی ← شعر + اسم فاعل + جمع + آمد + ه + بود + ند  
ب. صرف توزیعی ← شعرا + آم + ده + بو + د + ند

طبق تحلیل (الف)، سه واژ آمیخته در این جمله وجود دارد: «شعرا» که از عربی به فارسی راه یافته و «آمد» و «بود» که آمیزه‌ای از ریشه و تکواژ ماضی‌سازند. در تحلیل (ب)، واژه «شعرا» اساساً از داده‌ها کنار گذاشته شده است، زیرا ساختارهای زبان فارسی آن را تولید نکرده است. دو فعل «آمده» و «بودند» نیز واژ آمیخته نیستند؛ بلکه «آم» واژگونه ریشه «آ(ی)» است که در کنار وند صفت مفعولی «ده» قرار گرفته و «بو» نیز واژگونه ریشه «باش» است که با وند گذشته «د» همراه شده است. همچنین گفتنی است که نظریه صرف توزیعی، کسره اضافه تکواژ به‌شمار نمی‌رود؛ چراکه در چارچوب نظری یادشده برای کسره اضافه، جایگاه نحوی در نظر گرفته نمی‌شود. از این‌رو، کسره اضافه در این نظریه تنها به عنوان پیونددهنده آوایی شناخته می‌شود که می‌تواند میان صفت و موصوف (مانند: باغ زیبا)، مضاف و مضاف‌الیه (مانند: آموزگار من) و نیز در درون واژه‌ها قرار گیرد؛ مانند صفت «از دماغ فیل افتاده». باین‌همه، چه کسره اضافه را واژه‌بست تلقی کنیم، چه آن را پیونددهنده آوایی بدانیم، یا حتی ماهیتش را نشانگر حالت دستوری در نظر بگیریم، تغییری در تحلیل کمی تکواژها ایجاد نمی‌شود.

#### ۴ - ۲ - جامعه آماری پژوهش

اکنون به منظور تعیین رده ساختارهای زبان فارسی از دیدگاه صرف توزیعی، پیکره زبانی متنوعی از متون فارسی گردآوری می‌شود. برای دستیابی به نتیجه دقیق‌تر و صحیح‌تر

در این زمینه، متون یادشده به صورت همگن و متنوع انتخاب می‌شوند. گفتنی است که هدف از این پژوهش، تحلیل کلام و سبک‌شناسی متون فارسی نیست، با این‌همه، متونی که برگزیده می‌شوند، تنها به سبک، سیاق، حوزه و تخصص ویژه‌ای تعلق ندارند و کوشش می‌شود سبک‌های گوناگونی همچون سبک نوشتاری، سبک محاوره‌ای و سبک علمی و داستانی در پیکره گنجانده شود. همچنین، متون انتخابی حاوی مطالبی هستند که از حوزه‌های مختلف اجتماعی، فرهنگی، دین و مذهب، سیاسی، ورزشی، اقتصاد، پزشکی و سلامت، علوم مهندسی، زبان و ادبیات، معماری، تاریخ، صنعت و گردشگری برگزیده می‌شوند. از این‌رو، متون یادشده از منتخب سه کتاب داستان، دو کتاب درسی، یک فیلمنامه و نیز از مقاله‌های موجود در روزنامه‌ها و خبرگزاری‌های رسمی و سراسری ایران انتخاب می‌شوند. بدین ترتیب، در پیکره پژوهش تنوع سبک‌ها و حوزه‌های زبانی کمابیش رعایت می‌شود. پیکره مورد نظر، ۲۴ نوع متن گوناگون را تشکیل می‌دهد. از هر یک از ۲۴ نوع متن موجود، ۴۰۰ واژه به صورت تصادفی در قالب جمله نمونه‌گیری می‌شوند. شمار واژه‌های نمونه‌گیری‌شده از ۲۴ نوع متن یادشده، در مجموع ۹۶۰۰ واژه است که در روند پژوهش، بررسی و تحلیل می‌شوند. فهرست کتاب‌های درسی و کتاب‌های داستانی و مقالات و فیلمنامه در پیوست مقاله آمده است. در ادامه، به پیروی از گرینبرگ (۱۹۵۴) و نیز آیکنوالد (۲۰۰۷)، از شاخص کمی گرینبرگ برای تعیین میزان پیچیدگی درونی ساخت‌های دستوری واژه‌های زبان فارسی در پژوهش حاضر بهره گرفته شده است. گرینبرگ یادآور می‌شود که این شاخص به تعیین میزان ترکیب‌پذیری زبان‌ها اشاره دارد و از طریق تقسیم تعداد تکواژهای یک حجم نمونه، به تعداد واژه‌های موجود در آن به‌دست می‌آید. بر اساس آنچه در بخش مبانی نظری پژوهش آمد، در معیار کمی گرینبرگ زبان‌ها به سه دسته تحلیلی و ترکیبی و بس‌ترکیبی تقسیم می‌شوند. بدیهی است هرچه حاصل تقسیم به حداکثر محدوده نزدیک‌تر باشد، تمایل زبان برای داشتن ویژگی‌های رده بعد از خود افزایش می‌یابد (گرینبرگ، ۱۹۵۴: ۱۹۴). نتایج حاصل از بررسی میزان پیچیدگی درونی ساخت‌های دستوری واژه‌های فارسی که از تقطیع و تجزیه ساخت‌های صرفی نحوی واژه‌های متون منتخب بر اساس شاخص کمی گرینبرگ به‌دست آمده، در جدول زیر گرد آمده است.

جدول ۱. نتایج بررسی میزان پیچیدگی درونی ساخت‌های دستوری واژه‌ها در زبان فارسی

۱۶۶ ساختار و تحلیل پیوندی زبان فارسی بر پایه نظریه صرف توزیعی

تعداد واژه‌ها	تعداد تکواژها	تعداد میانگین	تعداد تکواژها در میانگین	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰
۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰
۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲	۱,۵۱۲

تعداد واژه‌ها	تعداد تکواژها	تعداد میانگین	تعداد تکواژها در میانگین	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰
۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰	۴۰۰
۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱	۱,۵۲۱

همان گونه که پیش تر نیز اشاره شد، متون برگزیده به حوزه‌های تخصصی و سبک‌های کلامی متنوع و گوناگونی تعلق دارند تا از رهگذر چنین انتخاب همگنی بتوان واژه‌های فارسی را در حوزه‌های مختلفی بررسی کرد و پیکره وسیع‌تر و متنوع‌تری را پوشش داد. یکسان بودن حجم نمونه‌ها یا همان تعداد واژه‌ها در متن‌ها نیز نشان‌دهنده تمایز بهتر تعداد تکواژها در نمونه‌ها خواهد بود. آن چنان که در جدول مشاهده می‌شود، از بررسی ۲۴ متن انتخابی که در مجموع ۹۶۰۰ واژه را در برمی‌گیرد، درمی‌یابیم که شمار تکواژها در حجم نمونه‌های یکسان، بسیار به یکدیگر نزدیک هستند و در همه متون، تعداد تکواژها در بازه عددی بین ۵۶۹ تکواژ (کمترین تعداد تکواژ) و ۶۳۱ تکواژ (بیشترین تعداد تکواژ) قرار می‌گیرند.

اکنون می‌توان نتیجه حاصل از به‌کارگیری شاخص کمی گرینبرگ را در پیکره مورد نظر که از طریق تقسیم تعداد تکواژهای یک حجم نمونه، به تعداد واژه‌های موجود در آن به دست می‌آید، مشاهده کرد. مطابق با جدول، اعداد مربوط به ردیف «میانگین تعداد تکواژها در متن» نشان می‌دهند اعدادی که از این شاخص به دست آمده‌اند، از ۱,۴۲۲ تا ۱,۵۷۷ متغیر بوده‌اند. نزدیک بودن اعداد به دست آمده از تقطیع تکواژی در هریک از متون، نشان از صحیح بودن تحلیل پژوهش دارد و قطعیت نتایج را در تقطیع تکواژی واژه‌های زبان فارسی و تعیین رده زبانی آن تأیید می‌کند. در نهایت، برای تعیین مقدار کمی کلی و تعیین رده زبانی زبان فارسی، میان مقادیر کمی‌ای که برای هریک از متون منتخب به دست آمده است، عمل میانگین‌گیری صورت می‌گیرد. مطابق با آنچه در جدول آمده، میانگین یادشده، عدد ۱,۵۰۶ است. مقدار کمی ۱,۵۰۶ بیانگر این است که بر اساس شاخص کمی گرینبرگ که عنوان می‌کند اگر زبانی بین ۱ تا ۱,۹۹ صدم تکواژ در هر واژه داشته باشد، تحلیلی است، زبان فارسی با عدد ۱,۵۰۶ به رده زبان‌های تحلیلی گرایش بیشتری دارد. از این رو، گرایش غالب ساخت‌های صرفی نحوی زبان فارسی به لحاظ رده‌شناسی ساختواژی به سوی رده زبان‌های تحلیلی متمایل است. همچنین گفتنی است از آنجا که حاصل تقسیم و عدد به دست آمده از این شاخص به حداکثر محدوده

نزدیک‌تر نیست، می‌توان اذعان کرد که زبان فارسی برای داشتن ویژگی‌های رده‌ بعد از خود، یعنی رده‌ زبان‌های ترکیبی، تمایل چندانی ندارد.

در ادامه پژوهش، بررسی واژه‌های فارسی را در محور عمودی شکل (۱) که مربوط به شفافیت مرز تکواژها یا همان راه‌های اتصال تکواژهاست، پی می‌گیریم. آنچه از بررسی‌ها حاصل می‌آید، به نظر می‌رسد که در محور عمودی، زبان فارسی به سوی زبان‌های پیوندی گرایش بیشتری دارد. هرچند برخی از واژه‌های این زبان از یک تکواژ ساخته می‌شوند و رده‌ی از زبان‌های گسسته در آن برجای می‌گذارند، یا شمار اندکی از واژه‌های آن به صورت هم‌آمیخته ساخته می‌شوند، اما اشتقاق و تصریف فرایندهای ساختوازی زبانی در فارسی‌اند و واژه‌هایی که از این رهگذر شکل می‌گیرند، عموماً بسیارند و دارای مرزهای روشنی میان تکواژهای سازنده‌شان هستند. در ویژگی‌ای که به زبان‌های پیوندی نسبت داده می‌شود نیز می‌بینیم که هر واژه می‌تواند از چندین تکواژ تشکیل شود، در حالی که مرز میان تکواژها روشن و مشخص است و به ندرت تکواژ آمیخته‌ای در این زبان‌ها یافت می‌شود (آیکنوالد، ۲۰۰۷: ۴). با بیان این مطالب و وجود واژه‌های بسیاری در زبان فارسی که به این طریق ساخته می‌شوند، به نظر می‌رسد می‌توان چنین ادعا کرد که در ساختوازی زبان فارسی، راه‌های اتصال تکواژها عمدتاً از الگوی زبان‌های پیوندی پیروی می‌کند. در فارسی اغلب وندهای تصریفی و اشتقاقی تنها حامل یک معنای دستوری یا واژگانی هستند و با حفظ مرزهای خود به ریشه متصل می‌شوند. مانند:

- (۱۳). الف. «پژوهشگران‌شان» از اتصال تکواژهای (پژوه + ش + گر + ان + شان) ساخته می‌شود.  
ب. «نمی‌خریدند» از اتصال تکواژهای (ن + می + خرید + د + ند) ساخته می‌شود.  
پ. «می‌سوزاند» از اتصال تکواژهای (می + سوز + ان + د + 0) ساخته می‌شود.  
ت. «درس‌خوانده‌تر» از اتصال تکواژهای (درس + خوان + ده + تر) ساخته می‌شود.  
ث. «قیمت‌گذاری» از اتصال تکواژهای (قیمت + گذار + ی) ساخته می‌شود.  
ج. «ورزشگاه‌ها» از اتصال تکواژهای (ورز + ش + گاه + ها) ساخته می‌شود.

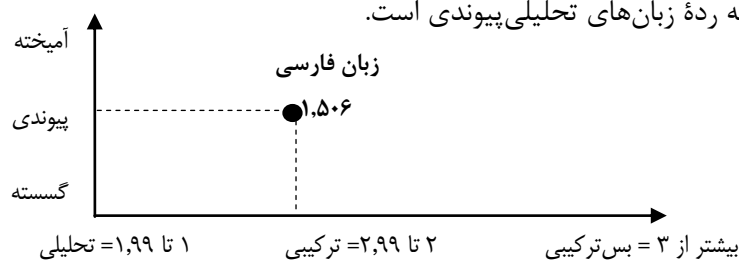
به این ترتیب، نتایج در بررسی تعیین میزان شفافیت مرز تکواژهای واژه‌های فارسی در پیکره پژوهش، حاکی از آن است که از میان ۹۶۰۰ واژه منتخب در پیکره، ۶۴ درصد از واژه‌ها، از طریق پیوندی و وندافزایی در دستگاه ساختوازی زبان فارسی ساخته می‌شوند و هر تکواژ در آنها صرفاً حامل یک معناست. همچنین گفتنی است، ۲۷ درصد از واژه‌های پیکره نیز به رده‌ی زبان‌های گسسته تعلق دارند که تناظر یک به یک میان تکواژها و واژه‌ها در آنها وجود دارد و هر تکواژ برابر با یک واژه است. یافته دیگر این بررسی نشان می‌دهد که ۹ درصد از واژه‌های پیکره در گروه زبان‌های آمیخته می‌گنجد که برخلاف زبان‌های پیوندی، مرز میان تکواژها در آنها مشخص نیست و عناصر سازنده واژه

را نمی توان به سادگی از هم تفکیک کرد. با توجه به نتایجی که از این بررسی حاصل می آید، می توان اذعان کرد که گرایش غالب رده ساختواژی زبان فارسی بر اساس معیار شفافیت مرز تکواژها، به سوی رده زبان های پیوندی است و اکثر واژه های فارسی با وندافزایی ساخته می شوند و معمولاً هر واژه فارسی در بردارنده چند تکواژ است که مرز میان آنها شفاف و روشن است.

جدول ۲. نتایج بررسی میزان شفافیت مرز تکواژها در واژه های فارسی

میزان	
زبان های گسسته	٪۲۷
زبان های پیوندی	٪۶۴
زبان های آمیخته	٪۹

در ادامه، یافته های به دست آمده از بررسی ها در مورد تعیین رده ساختواژی زبان فارسی، در محور دوبعدی ای که آیکنوالد (۲۰۰۷) از رده ساختواژی زبان ها ارائه می دهد، در شکل (۲) نمایش داده شده است. آنچنان که در شکل نیز دیده می شود، رده ساختواژی زبان فارسی بر روی محور در موقعیت [تحلیلی و پیوندی] قرار می گیرد. قرارگیری زبان فارسی در چنین مختصاتی در نمودار نشان دهنده گرایش غالب زبان فارسی به رده زبان های تحلیلی پیوندی است.



شکل (۲). رده ساختواژی زبان فارسی

## ۵- نتیجه

در مقاله حاضر کوشش شد بر پایه یافته های به دست آمده از انگاره دستوری صرف توزیعی، تحلیل و درونداد مناسبی برای تعیین رده ساختواژی زبان فارسی فراهم آید. بدین ترتیب، بر پایه این یافته ها، گرایش غالب رده ساختواژی زبان فارسی در نمودار دوبعدی ای که آیکنوالد (۲۰۰۷) از رده بندی صرفی زبان ها به دست داده و آن را بر اساس معیارهای ارائه شده ساپیر (۱۹۲۱) تنظیم کرده است، تعیین شد. در محور افقی نمودار یادشده که به میزان پیچیدگی درونی ساخت های دستوری واژه ها می پردازد، به پیروی از آیکنوالد و شاخص کمی گرینبرگ (۱۹۵۴)، با تقسیم تعداد تکواژهای هر جمله بر تعداد واژه ها، میزان تعلق رده ساختواژی زبان فارسی به هر یک از زبان های



تحلیلی و ترکیبی و بس ترکیبی مشخص شد. با بررسی ساخت‌های دستوری زبان فارسی در متونی که بدین منظور برگزیده شده بودند، نتیجه گرفته می‌شود که با به‌دست آمدن عدد ۱،۵۰۶ از حاصل تقسیم تعداد تکواژهای هر جمله بر تعداد واژه‌ها می‌توان اذعان کرد که زبان فارسی معاصر به لحاظ ترکیب‌پذیری واژه‌ها به رده‌ی زبان‌های تحلیلی گرایش بیشتری دارد. به این ترتیب، پرسش نخست پژوهش که به جایگاه زبان فارسی بر روی محور پیچیدگی درون ساخت‌های دستوری اشاره دارد، پاسخ داده می‌شود و زبان فارسی در این بررسی به رده‌ی زبان‌های تحلیلی تعلق می‌یابد و انگاره‌ی دستوری صرف توزیعی نیز همچون رهیافت صرف واژگرا، رده‌ی ساختواژی زبان فارسی را جزو زبان‌های تحلیلی قرار می‌دهد. همچنین یافته‌ی دیگر پژوهش را می‌توان چنین برشمرد که بر اساس معیار شفافیت مرز تکواژها، اغلب واژه‌های زبان فارسی، یعنی ۶۴ درصد از کل واژه‌های پیکره، از طریق پیوندی و وندافزایی که هر تکواژ صرفاً حامل یک معناست، در دستگاه ساختواژی زبان فارسی ساخته می‌شوند. از این‌رو، زبان فارسی به رده‌ی زبان‌های پیوندی گرایش بیشتری دارد، هرچند که ممکن است رده‌های دیگری از نوع گسسته یا هم‌آمیخته نیز در آن دیده شود. بدین ترتیب، می‌توان بر اساس معیار یادشده پرسش دوم پژوهش را که از منظر شفافیت مرز میان تکواژها، رده‌ی ساختواژی زبان فارسی را مورد پرسش قرار می‌دهد، پاسخ گفت. سرانجام این‌که می‌توان اذعان کرد که انگاره‌ی دستوری صرف توزیعی با وجود تفاوت‌هایی که در نوع و شمار تکواژها با دیگر انگاره‌های صرفی دارد، همچون انگاره‌های صرف واژه‌گرا، گرایش غالب رده‌ی ساختواژی زبان فارسی را به لحاظ ترکیب‌پذیری واژه‌ها و نیز شفافیت مرز تکواژها، در رده‌ی زبان‌های تحلیلی پیوندی قرار می‌دهد. بدین ترتیب، انگاره‌ی ناواژه‌گرای صرف توزیعی تأییدکننده‌ی تحلیلی پیوندی بودن رده‌ی ساختواژی زبان فارسی است که مؤید نتایج پژوهش‌های پیشین در صرف واژه‌گرا در این زمینه است. نکته‌ی قابل تألم دیگر این‌که در پژوهش‌های پیشین در این زمینه، همچون پژوهش شریفی و اخلاقی (۱۳۹۲) که بر انگاره‌ی صرف واژه‌گرا استوارند، اختلاف اندکی میان عدد حاصل از معیار گرینبرگ با عدد کمی به‌دست‌آمده در پژوهش حاضر، به چشم می‌خورد. نزدیک بودن نتایج در دو پژوهش را می‌توان برآیند دو اختلاف روش در این دو پژوهش دانست؛ اول، کنار نگذاشتن و شمارش تکواژهای وام‌واژه‌هایی که حاصل تولید دستگاه ساختواژی فارسی نیستند. دوم، تفاوت تعداد تکواژها در تقطیع تکواژی واژه‌ها در دو رویکرد صرف واژه‌گرا و صرف ناواژه‌گرا. همچنین این‌که می‌توان اذعان کرد که در پژوهش حاضر نزدیک‌تر بودن عدد ۱،۵۰۶ در بازه [۲ و ۱]، به عدد یک، نسبت به عدد ۱،۵۳ در پژوهش شریفی و اخلاق

(۱۳۹۲) حاکی از آن است که از منظر صرف توزیعی ساخت‌های صرفی نحوی زبان فارسی نسبت به پژوهش‌های پیشین، گرایش بیشتری به تحلیلی‌تر شدن دارند و به سوی تحلیلی‌تر شدن پیش می‌روند، و نه ترکیبی شدن.

### منابع

- آل احمد، جلال، (۱۳۴۵). مدیر مدرسه، سازمان کتاب‌های پرستو، چاپ سوم.  
 انوشه، مزدک، (در دست چاپ). صرف در نحو: از کمینه‌گرایی تا صرف توزیعی، تهران، دانشگاه تهران.  
 انوشه، مزدک، (۱۳۹۷). «بازنگری در تصریف زمان گذشته در زبان فارسی بر پایه نظریه صرف توزیعی»،  
 جستارهای زبانی، د ۹، ش ۱ (پیاپی ۴۳)، صص ۵۷-۸۰.  
 بیضایی، بهرام، (۱۳۸۰). سگ‌کشی، نشر روشنگران و مطالعات زنان، چاپ اول.  
 پایگاه دادگان زبان فارسی، پژوهشگاه علوم انسانی و مطالعات فرهنگی.  
 پرسپولیس نیوز، هفتم شهریور (۱۳۹۸). «اطلاعیه باشگاه پرسپولیس: هواداران برای حمایت از پرسپولیس به ورزشگاه بیایند».  
 ثمره، یدالله، (۱۳۶۹). «تحلیلی بر رده‌شناسی زبان: ویژگی‌های رده‌شناختی زبان فارسی»، مجله زبان‌شناسی،  
 س ۷، ش ۱، صص ۶۱-۸۰.  
 خبرگزاری تسنیم، دهم شهریور (۱۳۹۸). «مراسم رونمایی از پیراهن تیم اسکواش بانوان مهرسان با حضور  
 مسئولان فدراسیون برگزار شد».  
 خبرگزاری فارس، دهم شهریور (۱۳۹۸). «دانشگاه‌ها به سمت تغییر سرفصل‌های درسی پیش می‌روند» و  
 «جلسه نمایندگان آموزشگاه‌های آزاد هنری» و «بازدید سرپرست دفتر آموزش و توسعه فعالیت‌های  
 هنری از روند برگزاری آزمون‌های جامع و ادواری».  
 خبرگزاری مهر، دهم شهریور (۱۳۹۸). «بی‌خوابی با افزایش ریسک بیماری قلبی و سکنه مرتبط است».  
 دولت‌آبادی، محمود، (۱۳۶۹). روزگار سپری‌شده مردم سالخورده، کتاب اول، اقلیم باد، نشر چشمه، تهران،  
 چاپ اول.  
 سازمان پژوهش و برنامه‌ریزی آموزشی، (۱۳۹۸). تاریخ ۳، ایران و جهان معاصر، پایه دوازدهم، دوره دوم  
 متوسطه، رشته علوم انسانی، شرکت چاپ و نشر کتاب‌های درسی ایران، چاپ دوم.  
 سازمان پژوهش و برنامه‌ریزی آموزشی، (۱۳۹۶). زیست‌شناسی ۱، پایه دهم، دوره دوم متوسطه، رشته علوم  
 تجربی، شرکت چاپ و نشر کتاب‌های درسی ایران، چاپ دوم.  
 شریفی، شهلا. الهام اخلاقی، (۱۳۹۲). «بررسی ساختار و زبان فارسی امروز بر پایه رویکرد کمی رده‌شناسی»،  
 هشتمین همایش بین‌المللی انجمن ترویج و زبان و ادب فارسی ایران، دانشگاه زنجان، ۱۳ تا ۱۵ شهریور.  
 دبیرمقدم، محمد، (۱۳۹۲). رده‌شناسی زبان‌های ایرانی، جلد اول، سمت، چاپ اول.  
 رضایی، والی، فاطمه بهرامی، (۱۳۹۴). مبانی رده‌شناسی زبان، دانشگاه شهید بهشتی، چاپ اول.  
 روزنامه ابتکار، دهم شهریور (۱۳۹۸). «مستاجران از اینجا مانده، از آنجا رانده».  
 روزنامه جوان، دهم شهریور (۱۳۹۸). «احتمال توقف تولید پراید و ۴۰۵».  
 روزنامه جمهوری اسلامی، دهم شهریور (۱۳۹۸). «برای ماندگاری مکتب عاشورا».  
 روزنامه خراسان، دهم شهریور (۱۳۹۸). «بورس ۱۱ میلیون نفری شد» و «چالش‌های پس از جدایی» و  
 «مسمومیت ۳۳۲ نفر در شاندریز».  
 روزنامه کیهان، دهم شهریور (۱۳۹۸). «ماهواره‌های آمریکا را چه کسی منهدم کرد؟» و «سوداگری در بازار  
 مسکن نتیجه انفعال در اخذ مالیات از خانه‌های خالی».  
 روزنامه مردم‌سالاری، دهم شهریور (۱۳۹۸). «رایزنی با مجامع هواداران برای حضور زنان در ورزشگاه‌ها و  
 تماشای لیگ فوتبال» و «آغاز عملیات بازار باز در شهریورماه».

- روزنامه هم‌ملکی، دهم شهریور (۱۳۹۸). «بعد از مرمت و احیای خانه جلال، این بار خانه همسایه او یعنی نیما یوشیج چشم‌انتظار بازسازی است» و «تعریفی از مردم».
- هدایت، صادق، (۱۳۲۱). سگ ولگرد: تاریخ‌خانه، امیرکبیر، چاپ پنجم، تهران.
- Aikhenvald, A.Y., 2007. Typological Distinctions in Word-Formation. In *Language Typology and Syntactic Description Vol III: Grammatical Categories and the Lexicon*, T. Shopen (Ed.), 2<sup>nd</sup> ed., Cambridge, Cambridge University Press, p 1-65.
- Al-e-Ahmad, J., 1966. *Headmaster (modir-e madrese)*, sazman-e ketabha-ye parastu, Third Edition. [In Persian].
- Anushe, Mazdak, (in press). *Morphology in Syntax: From Minimalist Program to Distributed Morphology*, Tehran: University of Tehran. [In Persian].
- Anushe, M., 2018. "A Revision of Persian Past Tense Inflection a Distributed Morphology Approach", *Language Related Research*, volume 9, issue 1. P. 57-80. [In Persian].
- Bayzai, B., 2001. *Killing Mad Dogs (Sagkoshi)*, Roshangaran and Women Studies publication, First Edition. [In Persian].
- Bloomfield, L., 1933. *Language*, New York: Henry Holt.
- Chomsky, N., 1995. *The Minimalist program*, MIT Press, Cambridge.
- Comrie, B., 2001. Different Views of Language Typology. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language Typology and Language Universals* (pp. 25-39). Berlin & New York: Walter de Gruyter.
- Dabir Moghaddam, M., 2013. *Typology of Iranian Languages*, Volume 1, Tehran: SAMT, First Edition. [In Persian].
- Dolat Abadi, M., 1990. *Time spent on older people; Book I: (climate, wind)*, Tehran: Cheshmeh Publishing, First Edition. [In Persian].
- Ebtekar* Newspaper, 1 September 2019. [In Persian].
- Educational Research and Planning Organization, 2019. *History 3: Iran and the contemporary world*, Twelfth grade, Secondary high school, Field of Humanities, Iran Textbook Publishing Company, second edition. [In Persian].
- Educational Research and Planning Organization, 2017. *Biology 1*, tenth grade, Secondary high school, Field of Experimental Sciences, Iran Textbook Publishing Company, second edition. [In Persian].
- Embick, D., 2015. *The Morpheme*, Berlin: Mouton de Gruyter.
- Fars* News Agency, 1 September 2019. [In Persian].
- Greenberg, J.H., 1954. "A Quantitative Approach to the morphological Typology of Language", in Spencer, R. (Ed), *Method and Perspective in Anthropology*, Minneapolis.
- Greenberg, J.H., 1974. *Language Typology: A Historical and Analytic Overview*. The Hague, Paris: Mouton.
- Halle, M. & A. Marantz., 1993. Distributed Morphology and the pieces of inflection. In: K. Hale & S.J. Keyser (Eds). *The View from Building 20. Essays in Linguistics in Honour of Sylvain Bromberger*, Cambridge, MA: MIT Press, p111-176.
- Hamdeli* Newspaper, 1 September 2019. [In Persian].
- Hedayat, S., 1942. *The Stray Dog (Sag-e velgard): Darkroom (tarikkhane)*, Tehran: Amir Kabir, Fifth Edition. [In Persian].

- Hockett, Ch., 1958. Two models of grammatical description. In: M. Joos (Ed). *Readings in linguistics* [2<sup>nd</sup> Ed], Chicago: University of Chicago Press, p386-399.
- Humboldt, W. v., 1825. Über das Entstehen der grammatischen Formen, und ihren Einfluss auf die Ideeentwicklung. *Abhandlungen der hist. phil. Klasse*, Berlin: königliche Akademi der Wissenschaften, 401-30.
- Humboldt, W. v., 1836. *Über die Verschiedenheit des menschlichen Sprachbaus (Vol. I of Über die Kawispache auf der Insel Java)*. Berlin: königliche Akademi der Wissenschaften.
- Javan Newspaper*, 1 September 2019. [In Persian].
- Jomhour-e Eslam Newspaper*, 1 September 2019. [In Persian].
- Kayhan Newspaper*, 1 September 2019. [In Persian].
- Khorasan Newspaper*, 1 September 2019. [In Persian].
- Marantz, A., 1997. No Escape from Syntax; Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In: A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams (Eds). *University of Pennsylvania Working Papers in Linguistics*, Philadelphia: University of Pennsylvania, 201-225.
- Mardomsalari Newspaper*, 1 September 2019. [In Persian].
- Mehr News Agency*, 1 September 2019. [In Persian].
- Nida, E., 1949. *Morphology: The Descriptive Analysis of Words*, Ann Arbor: The University of Michigan Press.
- Nida, E., 1965. *Morphology: The Descriptive Analysis of Words* [2<sup>nd</sup> Ed], Ann Arbor: University of Michigan Press (First published 1949).
- Noyer, R., 1997. *Features, Positions and Affixes in Autonomous Morphological Structure*, NY: Garland.
- Persian Linguistic Database (PLDB), Institute for Humanities and Cultural Studies. [In Persian].
- Perspolis News Agency*, 29 August 2019. [In Persian].
- Rezaei, vali. Bahrami, F., 2015. *Fundamentals of Linguistic typology*, Tehran: Shahid Beheshti University, First Edition. [In Persian].
- Samareh, Y., 1990. "An analysis of the typology of language: typology features of Persian language", *Iranian journal of linguistics*, volume 7, issue 1. P. 61-80. [In Persian].
- Sapir, E., 1921. *Language*, New York: Harcourt, Brace, and World.
- Schlegel, A.W.v., 1818. *Observations sur la langue et la littérature provinciales*. Parise.
- Schlegel, K. W. F. v., 1808. *Über die Sprache und Weisheit der Indier*. Heidelberg: Mohr and Zimmer.
- Sharifi, Shahla. Akhlaghi, e., 2013. "A study of the Morphology of Persian Language Today Based on a Quantitative Approach to Typology", The 8th International Conference of the Association for the Promotion of Persian Language and Literature of Iran, University of Zanjan. [In Persian].
- Shopen, T., 2007, *Language Typology and Syntactic Description*, Volume 3: Complex Constructions (2nd Edition), Cambridge University Press.
- Siddiqi, D., 2009. *Syntax within the Word*. Amsterdam: John Benjamins Publishing.
- Tasnim News Agency*, 1 September 2019. [In Persian].
- Thompson, L. C., 1987. *A Vietnamese Reference Grammar*. Honolulu: University of Hawai'i Press.