کتابهای درسی زبان‌شناسی رایانه‌ای:
زبانهای برنامه نویسی(۱)

الهه دفتری نژاد
daftarynejad@yahoo.com

چکیده

زبان‌شناسی رایانه‌ای شاخه‌ای میان رشته‌ای است. اگر مطالعات زبان‌شناسی رایانه‌ای را به دو گروه کل‌گرا(گرایشهای شناختی) و کمینه‌گرا (گرایشهای مبتنی بر احتمالات) تقسیم کنیم، می‌توان گفت که در زبان‌شناسی رایانه‌ای معمولاً گرایشهای شناختی با زبانهای رایانه‌ای اعلانی(برای مثال پرولوگ) کدنویسی می‌شود.اما مطالعات آماری به کاربرد زبانهای برنامه‌نویسی رویه‌ای (برای مثال، سی و پاسکال) گرایش دارند؛ زیرا این زبانها به صورت استانده شامل توابع ریاضی اند که با برنامه نویسی محاسبات آماری هم‌سوتر است. با توجه به مباحثی که اکنون در زمینه زبان‌شناسی رایانه‌ای در دنیا مطرح است ــ منظور گسترش روش شناسی پیکره، ضرورت تحلیلهای مبتنی بر اصول آمار و احتمالات و سایر تحلیلهای ریاضی بنیان در مطالعه زبان و مانند آنها ــ می‌توان گفت برای پژوهشگران زبان‌شناسی رایانه‌ای، دانستن یکی از زبانهای برنامه‌نویسی رویه‌ای کاربردی‌تر است.

واژه‌های کلیدی

زبان‌شناسی‌رایانه‌ای، پردازش زبان طبیعی، برنامه نویسی، رایانه.

مقدمه

زبان‌شناسی رایانه‌ای[۱] شاخه‌ای میان رشته‌ای است که در مراکز دانشگاهی دنیا به سرعت در حال گسترش است. از آنجا که در جهان،عرضه و تقاضا نقش تعیین‌کننده‌ای در جذب سرمایه‌های بخش خصوصی یا دولتی در مراکز آموزشی دارد، گسترش این رشته ناشی از تقاضای بازارهای جهانی است. در کشور ما نیز نه تنها این تقاضا وجود دارد، بلکه حتی نیازما به عنوان یک کشور در حال توسعه به متخصصان این رشته بسیار بیشتر است. برای مثال، بومی سازی دانش و اطلاعات موجود در شبکه جهانی اینترنت را در نظر بگیرید. این امر مهم که نقش مؤثری در بهره‌گیری همگانی از این ابزار جدید فناوری دارد، به پژوهشگران این رشته نیازمند است. پرواضح است که هم اکنون سهم اندکی از دانش و اطلاعات موجود در این شبکه به زبان فارسی است و ایجاد تغییر در این ساختار جهت حفظ و گسترش زبان فارسی و نیز بهره‌گیری مؤثر کاربران فارسی زبان کاملاً ضروری است. در عصر کنونی، گسترش زبان فارسی و حفظ آن در برابر وام‌گیری واژه‌های بیگانه، صرفاً با روشهای سنتی به سختی امکان‌پذیر است. حفظ و گسترش زبان فارسی، بومی سازی دانش جهانی و حتی انجام دادن پژوهشهای زبان شناختی و بسیاری از موارد مشابه دیگر، نیازمند بهره‌گیری از ابزارهای جدیدی چون پیکره‌های رایانه‌ای، فرهنگهای الکترونیک، فرهنگهای هوشمند و نرم افزارهای مختلف پردازش زبان است. تهیه، طراحی و نیز بهره‌مندی از این ابزارها خود در گرو تربیت نیروی انسانی متخصص در شاخهٔ زبان‌شناسی رایانه‌ای است.

با توجه به اهمیت کاربردی درس زبان‌شناسی رایانه‌ای، به جاست که به وضعیت ساعات آموزشی آن در ایران و سایر کشورها بپردازیم. با یک جستجوی ساده می‌توان دریافت که اکنون در بسیاری از دانشگاههای معتبر دنیا[۲]، درس زبان‌شناسی رایانه‌ای در کارشناسی ارشد به صورت تخصصی تدریس می‌شود. در برخی دیگر از دانشگاهها[۳]، واحدهای این درس به طور گسترده در گروه زبان‌شناسی همگانی ارائه می‌شود.[۴] در حالی که در کشور ما تنها دو واحد درسی به زبان‌شناسی رایانه‌ای اختصاص یافته است. این دو واحد جزو دروس انتخابی است و فقط در برخی از گروه‌های زبان‌شناسی همگانی تدریس می‌شود. با توجه به اینکه زبان‌شناسی همگانی در ایران در کارشناسی ارشد تدریس می‌شود، دانشجویان آن، باید ابتدا دروس تخصصی زبان‌شناسی و پیش‌نیازهای ضروری آن را بگذرانند؛[۵] در نتیجه، ساعاتی که برای مطالعه درس زبان‌شناسی رایانه‌ای و سایر دروس بین رشته‌ای زبان‌شناسی باقی می‌ماند به اجبار بسیار محدود می‌شود.

منابع درسی و پژوهشی رشتهٔ زبان‌شناسی رایانه‌ای عمدتاً زبان انگلیسی است. این امر نه تنها مانع از بومی شدن دانش این رشته در زبان فارسی می‌شود، بلکه مشکلات درک زبان بیگانه را به همراه دارد. این منابع دارای اصطلاحات تخصصی و کوته‌واژه‌های[۶] بسیاری است که بیشتر آنها معادل فارسی دقیق و پذیرفته‌ای ندارند. مسئلهٔ دیگر که مطالعه و درک این متون را تا حدی دشوار کرده است، ورود اصطلاحات رایانه‌ای در آنهاست. شاخهٔ زبان‌شناسی رایانه‌ای، چنان که از عنوان آن مشخص است، وابسته به علوم رایانه‌ای است. این موضوع باعث می‌شود که اصطلاحات رشتهٔ رایانه وارد متون زبان‌شناسی رایانه‌ای شود. ترکیب این اصطلاحات با اصطلاحات زبان‌شناسی به پیدایش اصطلاحات ترکیبی می‌انجامد که در کل برای دانش‌آموختگان زبان‌شناسی همگانی ناآشناست. ورود برخی فرمانهای[۷] برنامه نویسی و شبه برنامه نویسی در برخی از منابع این رشته از دیگر مسائلی است که باید بدان توجه کرد. برای مطالعهٔ این گونه منابع، نیاز به دانش و تجربهٔ قبلی در زمینهٔ برنامه نویسی است. در کل، عواملی که به آنها اشاره شد باعث می‌شود در گروههای زبان‌شناسی پژوهشهای کمتری در حوزه زبان‌شناسی رایانه‌ای صورت گیرد که در نتیجه این شاخهٔ پرکاربرد طرفداران اندکی خواهد داشت.

برای کاهش موانع موجود در مطالعه و درک منابع درسی و پژوهشی زبان‌شناسی رایانه‌ای، می‌توان راه کارهای متفاوتی ارائه کرد. برای مثال، آموزش مباحثی چون مبانی و مقدمات رایانه، برنامه‌نویسی و طراحی پایگاههای اطلاعاتی می‌تواند بسیار سودمند باشد. آموزش این گونه مباحث که معمولاً اصطلاحات تخصصی و کوته واژه‌های رشتهٔ رایانه‌ای دارد، درک متون زبان‌شناسی رایانه‌ای را بهبود می‌بخشد. مهم‌تر آن که پژوهش در این شاخه بدون داشتن دانش رایانه‌ای تقریباً ناممکن است؛ بنابراین، این علوم از جمله ضروریات این شاخه محسوب می‌شود. صحت این مطلب هنگامی آشکارتر می‌شود که واحدهای درسی رشتهٔ رایانه‌ای، به خصوص برنامه‌نویسی، جزو واحدهای آموزشی زبان‌شناسی رایانه‌ای دانشگاههای معتبر دنیا باشد.[۸] البته ممکن است برخی از پژوهشگران زبان‌شناسی رایانه‌ای، ضرورت یادگیری برنامه نویسی را با شک و تردید بنگرند. به نظر هاچینز[۹] دانستن حداقل یک زبان برنامه نویسی و آگاهی از روش طراحی و کار با پایگاههای اطلاعاتی برای پژوهشگران این رشته الزامی است(هاچینز، ۱۹۸۶). اگرچه برخی از متون این شاخه، به خصوص برخی از مباحث نظری، بدون آشنایی با علوم رایانه‌ای برای زبان‌شناسان قابل درک است، نمی‌توان این واقعیت را منکر شد که مطالعهٔ بسیاری ازمنابع این رشته اجرا و آزمودن نظریات و حتی درک اصولی مباحث نظری، به دانستن زبانهای

برنامه‌نویسی و طراحی پایگاههای اطلاعاتی وابسته است. برای مثال، برنامه نویسی حالت - متناهی[10] را در نرم افزار Xerox[11] در نظر بگیرید. کنت بیسلی[12] در مقدمهٔ کتاب *ساخت واژهٔ حالت - متناهی* در مورد این برنامه می‌گوید: اگرچه در اغلب کتابهای برنامه نویسی ادعا می‌شود که نیازی به دانش و تجربهٔ قبلی در زمینهٔ زبانهای برنامه نویسی نیست، من این مطلب را رد می‌کنم. برای مطالعه این کتاب به زبانهایی چون Java ،C++ و C باید از دانش برنامه‌نویسی برخوردار بود (بیسلی و همکار، ۲۰۰۳).

اگردانستن زبان برنامه نویسی برای محققان زبان‌شناسی رایانه‌ای الزامی است، حال کدام زبان برنامه‌نویسی مناسب‌تر است و آیا اصلاً از این دیدگاه تفاوتی میان زبانهای برنامه‌نویسی وجود دارد یا خیر؟ مکنری[13] در *زبان‌شناسی رایانه‌ای* تاحدودی به این مطلب پرداخته است(مکنری، ۱۹۹۲: ۲۰-۴۰) . وی در این اثر که از منابع درسی رشتهٔ زبان‌شناسی رایانه‌ای است، به پرسشهای مقدماتی این رشته پاسخ می‌دهد. *زبان‌شناسی رایانه‌ای* از حیث گزینش مباحث مقدماتی، ارتباط مفاهیم انتخابی با یکدیگر و نحوهٔ توضیح آنها اثری ارزشمند است. مکنری در مقدمهٔ کتاب چنین می گوید: «حیطهٔ وسیعی از جریان اصلی زبان‌شناسی رایانه‌ای، به عمد، دراین کتاب گنجانده نشده است؛ زیرا به اعتقاد من، آنچه در اینجا مطرح شده است، حداقل پیشینهٔ زبان‌شناسی و رایانه‌ای مورد نیاز برای پژوهشگران این رشته است» (همانجا).

در حقیقت چنین است. مکنری در این کتاب، به خصوص در فصلهای نخست آن، مطالبی را مطرح می کند که سایر نویسندگان معمولاً به دلایل مختلف از طرح آن گریزانند. کتاب *زبان‌شناسی رایانه‌ای* مکنری را می‌توان به دو بخش مشخص تقسیم کرد: بخش نخست که به توضیحات مقدماتی درباره زبان‌شناسی رایانه‌ای می‌پردازد و بخش دوم که برخی از طرحهای انجام شده در زبان انگلیسی را معرفی کرده آنها را شرح می‌دهد.

با توجه به موارد فوق ، نگارنده بر آن شد که بخش نخست کتاب *زبان‌شناسی رایانه‌ای* مکنری را خلاصه و ترجمه کند. این ترجمه با سه موضوع «معرفی زبان‌شناسی رایانه‌ای»، «معرفی حوزه‌های زبان‌شناسی رایانه‌ای» و «انتخاب زبانهای برنامه‌نویسی مناسب» در سه شمارهٔ پی در پی در فصلنامه *سخن سمت* منتشر خواهد شد. از آنجا که به نظر می‌رسد موضوع انتخاب «زبانهای برنامه‌نویسی مناسب» کاربرد بیشتری دارد، آن را پیش از مباحث مقدماتی‌تر بررسی می کنیم. به طور کلی، ترجمهٔ کتابهای درسی زبان‌شناسی رایانه‌ای از جمله راه کارهای مؤثر در تسهیل مطالعهٔ این منابع است، اما با دو مسئله مهم مواجه‌ایم: نخست، معادل گذاری اصطلاحات در زبان فارسی است. چنان که پیشتر آمد، بیشتر

اصطلاحات این شاخه فاقد معادلهای پذیرفته در زبان فارسی است. دوم آن که، گستردگی حوزه‌های این شاخه و تخصصی شدن هر یک از آنهاست. این گستردگی، معرفی جامع زبان‌شناسی رایانه‌ای را در گرو ترجمهٔ متون حوزه‌های مختلف این شاخه قرار می‌دهد؛ از جمله: فرهنگ نگاری، ترجمهٔ ماشینی، پردازش متنی و پردازش گفتاری. البته باید ترجمهٔ مطالب روش‌شناسی پیکره، روش طراحی پیکره‌های رایانه‌ای، شناخت انواع پیکره‌ها ومطالبی از این نوع را نیز به مجموعهٔ نخست افزود. با وجود این، اولویت‌بندی میزان اهمیت حوزه‌های گوناگون جهت معرفی مقدماتی آنها، از طریق ترجمه یا تألیف، سودمند است.

اینک به مباحث مقدماتی زبان‌شناسی رایانه‌ای می‌پردازیم و سپس با شرح دسته‌بندیهای بنیادین در زبان‌شناسی، که بیشتر زبان‌شناسان با آن آشنایی کامل دارند، آنها را در میان زبانهای برنامه‌نویسی بررسی می‌کنیم. در پایان، به بررسی ارتباط میان این دسته بندیها، مقایسهٔ آنها با موضوعات مطرح در زبان‌شناسی معاصر و نیز نتیجه‌گیری آن می‌پردازیم.[۱۴]

زبان‌شناسی رایانه‌ای چیست؟

ابتدا، باید دید زبان‌شناسی رایانه‌ای چیست؟ شاید پاسخ به این سؤال ساده به نظر برسد، اما چنین نیست. میتکو[۱۵] (۲۰۰۳) در کتاب *راهنمای زبان‌شناسی رایانه‌ای* که از برترین منابع درسی این شاخهٔ میان رشته‌ای است، در فهرست واژگان تعریف دقیقی از آن ارائه نمی‌کند و فقط در مقدمهٔ اثرش چنین تعریف می‌کند: «زبان‌شناسی رایانه‌ای شاخه‌ای میان رشته‌ای است که با استفاده از رایانه به پردازش زبان می‌پردازد. ... این شاخه به لحاظ نظری، با پیدایش الگوهای صوری و رایانه‌ای زبان گسترش یافت» (میتکو، ۲۰۰۳: IX). در بسیاری از منابع دیگر چنین تعریف جامع و مانعی از زبان‌شناسی رایانه‌ای مشاهده نمی‌شود؛ زیرا پرداختن به آنها خارج از حوزهٔ بحث کنونی است. جرفسکی[۱۶] با در نظر گرفتن سیر تحولی این رشته در دهه‌های گذشته و با دسته‌بندی‌ای کلی چنین می‌گوید: «از نظر تاریخی رشته‌های رایانه، مهندسی برق، زبان‌شناسی و روان‌شناسی/ دانش شناخت به گونه‌ای بسیار متفاوت به پردازش زبان و گفتار پرداخته‌اند. گوناگونی پردازش زبان و گفتار در میان این گروه‌های دانشگاهی مختلف موجب پدید آمدن رشته‌های متفاوتی شده است که در عین حال با یکدیگر همپوشی دارند. این رشته‌ها عبارت است از: زبان‌شناسی رایانه‌ای در زبان‌شناسی، پردازش زبان طبیعی[۱۷] در رایانه، تشخیص گفتار[۱۸] در مهندسی برق و روان زبان‌شناسی رایانه‌ای[۱۹] در روان‌شناسی (جرفسکی و همکار، ۲۰۰۰: ۱۰). این تقسیم‌بندی که

بیشتر متخصصان این رشته آن را تأیید کرده‌اند، حاصل چند دهه پژوهش در زمینهٔ پردازش زبان است؛ همچنین جایگاه این شاخه را در میان مجموعه رشته‌هایی که به موضوع پردازش زبان و گفتار می‌پردازند، به خوبی مشخص می‌کند. با توجه به این تعریف، به نظر می‌رسد هم‌پوشی میان رشته‌های فوق کاملاً ناهمگن است، بدین معنا که برخی از این رشته‌ها با یکدیگر هم پوشی و قرابت بیشتری دارند. بازتاب این مسئله در برخی از تعاریف به خوبی مشهود است. برای مثال، مکنری زبان‌شناسی رایانه‌ای را بدین صورت تعریف می‌کند: «زبان‌شناسی رایانه‌ای فصل مشترک میان دو رشتهٔ زبان‌شناسی همگانی و هوش مصنوعی[۲۰]، زیر شاخه‌ای از رشتهٔ رایانه است» (مکنری،۱۹۹۲:۱). این تعریف که تصویر روشن‌تری از زبان‌شناسی رایانه‌ای ارائه می‌دهد، تعریفی وابسته و غیرمستقل است. رشته‌های پدید آورندهٔ زبان‌شناسی رایانه‌ای ، یعنی زبان‌شناسی و هوش مصنوعی، به لحاظ تعاریف اولیه با دشواریها و گاهی با کاستیهایی مواجه‌اند که به طور مستقیم بر تعریف زبان‌شناسی رایانه‌ای تأثیر می‌گذارد. شرح این کاستیها که بیشتر متخصصان این رشته‌ها با آن آشنا نیستند، خارج از حوصلهٔ این بحث است. به همین سبب در ادامهٔ مطلب، بدون توجه و اشاره به این کاستیها و نیز بدون بررسی سایر تعاریف، تعریف اخیر را که ارتباط بیشتری با مطلب کنونی دارد مبنا قرار می‌دهیم.

گرایش شناختی[۲۱] و گرایش احتمالاتی[۲۲] در زبان‌شناسی رایانه‌ای

به طور کلی، می‌توان مطالعات زبان‌شناسی رایانه‌ای را در امتداد خطی فرضی ـ که پژوهشهای هوش مصنوعی را نیز دسته بندی می‌کند ـ در دو گروه مشخص تقسیم کرد. این خط فرضی، در واقع خط میان گرایشهای کل‌گرا[۲۳] و کمینه‌گرا[۲۴] نسبت به نظامهای هوشمند[۲۵] است. این تقسیم بندی نشان دهنده دسته‌بندی کلی در مطالعات زبان‌شناسی میان گرایشهای شناختی و احتمالاتی در توضیح فرایندهای زبانی است. چامسکی[۲۶] (۱۹۶۵) و شانون و وی ور[۲۷](۱۹۴۹)، به ترتیب، نمونه‌های شاخص هر یک از این دو گرایش‌اند. دو گرایش شناختی و احتمالاتی در زبان‌شناسی رایانه‌ای نیز اساساً با پدیدهٔ زبان به دو شیوهٔ متفاوت برخورد می‌کنند. معمولاً مطالعات شناختی، کل‌گرا و مطالعات احتمالاتی کمینه‌گرا هستند. بنابراین، خط فرضی میان کل‌گرایی[۲۸] و کمینه‌گرایی[۲۹]، به پژوهشهای زبان‌شناسی، هوش مصنوعی و نیز زبان‌شناسی رایانه‌ای تقسیم می‌شود.

کل‌گرایی پدیده‌هایی را بررسی می‌کند که پیامد جانبی یک نظام‌اند. این بررسی از دیدگاه نظامی که آن پیامد را ایجاد کرده است انجام می‌شود، نه از دیدگاه ویژگیهای

پیامد. از سوی دیگر، کمینه‌گرایی نظامها را به عنوان مجموعه پدیده‌هایی در نظر می‌گیرد که در واقع محصول آن نظام هستند. می‌توان گفت، دیدگاه کل‌گرا می‌کوشد تا پدیده‌های گوناگون مرتبط با هر نظام را به نوعی به علت مشترکی مربوط سازد تا تصویری کلی از نظام اصلی که سبب وقوع آن پدیده‌ها شده است به دست آورد. اما دیدگاه کمینه‌گرا می‌کوشد تا هر نظامی را با توضیح و برشمردن پدیده‌هایی که آن نظام به وجود آورده است توضیح دهد. در نتیجه، در نگرش کل‌گرا سعی بر آن است که هر نظامی به صورت یک واحد کل نمونه‌سازی شود، در حالی که در نگرش کمینه‌گرا سعی می‌شود تا با نمونه‌سازی تک تک پدیده‌های مرتبط با آن نظام، در نهایت کل نظام نمونه سازی شود.

نگرشهای کل‌گرا و کمینه‌گرا در پدیدهٔ زبان

کل‌گرایی نسبت به پدیده زبان یکی از ویژگیهای بینش سوسوری است. سوسور اجزاء زبان را به عنوان عوامل تشکیل دهندهٔ یک «کل» می‌پنداشت که هر کدام جایگاه و نقشی در درون آن «کل» به عهده داشت. وی عقیده داشت که آن «کل» یعنی «زبان» خود پدیده‌ای ذهنی و روان‌شناختی است ... (دبیر مقدم،۱۵:۱۳۷۸). از دیدگاه کل‌گرایی، زبان یکی از پدیده‌های ناشی از مغز است؛ بنابراین برای نمونه سازی زبان باید نظام اصلی پدیدآورندهٔ آن یعنی مغز نمونه‌سازی شود. در نتیجه، دیدگاه کل‌گرا نسبت به زبان درواقع پیرو رویکرد شناختی است. علم شناخت می‌کوشد تا وضعیت مغز را هنگام انجام دادن فرایندهای استدلال و ادراک نمونه‌سازی کند و نشان دهد که ذهن انسان برای درک جهان به چه تمهیداتی اندیشیده است. آشکار است که مطالعهٔ علوم شناختی، نقطهٔ آغاز مطالعات کل‌گرا نسبت به فرایند زبان است، زیرا مطالعات شناختی می‌تواند واقعیتهای ساختار ذهن انسان را بیان کند؛ ساختاری که مطالعات کل‌گرای پردازش، زبان طبیعی در تلاش برای مشابه‌سازی آن هستند. پژوهشگرانی که تمهیدات ذهنی انسان را در ماشین پیاده می‌کنند، معتقدند که اگر گرایش کل‌گرا مؤثر باشد، باید پس از نمونه‌سازی «کل» نظام مغز درماشین، تمامی پدیده‌های وابسته به آن درماشین بروز کند.

دیدگاه کل‌گرا نسبت به فرایند زبان دارای برخی فواید کاربردی است. یکی از فواید آن تعیین ملاک مشترک برای پژوهشگرانی است که سایر پدیده‌های وابسته به مغز را بررسی می‌کنند. دیدگاه کل‌گرا می‌تواند موجب آسان‌سازی تبادل اطلاعات میان پژوهشگران این رشته‌ها و نیز مانع از تلاشهای تکراری آنان باشد. دیدگاه کل‌گرا کاستیهایی نیز دارد و آن هنگامی بروز می‌کند که نتوان بنیادهای نظام هدف را نمونه‌سازی کرد یا هنگامی که

نمونه‌سازی نظامی به زمان طولانی نیاز داشته باشد، اما نتایج آن هم اکنون مورد نیاز باشد. به عبارتی دیگر، به هیچ وجه مشخص نیست که آیا می‌توان فرایندهای شناختی مرتبط با مغز را در رایانه نمونه‌سازی کرد و آیا اصلاً این نمونه‌سازی در حوزهٔ علوم رایانه کاربردی دارد یا خیر؟ اگر نتوان فرایندهای شناختی را به وسیله رایانه مشابه‌سازی کرد، نمی‌توان نظام آن، یعنی مغز را نمونه‌سازی کرد و اگر نتوان نظام را با فرض پیروی از راه کار کل‌گرا نمونه سازی کرد نمی‌توان به وسیله رایانه پدیده‌های مغزی ایجاد کرد. همچنین در مشابه‌سازی نظام مغز، به صورت کامل، زمان زیادی صرف می‌شود و از طرفی رسیدن به نتایج آن فوریت دارد؛ یعنی زودتر از حداقل زمانی که برای حصول این نتایج بر پایهٔ اصول کل‌گرایی لازم است. یکی از شاخه‌های گرایش شناختی که درواقع کمینه‌گرا شده است، در قالب طراحی «رابطهای زبان طبیعی»[۳۰] بررسی می‌شود. در این شاخه سعی بر آن است تا از نظامهای مورد تقاضا، تعبیری جدید ارائه شود و سپس نمونه‌سازی صورت گیرد. در این تعبیر جدید، زیر مجموعهٔ خاصی از نظام مورد نظر که معمولاً مرتبط با یک زمینهٔ کاربردی است، به عنوان نظام کامل تعریف می‌شود. با وجود این، نمونه‌سازیهای شناختی زبان عموماً پیرو تفکر کل‌گرایی‌اند. اگرچه می‌توان «پایانه‌های زبان طبیعی»[۳۱] را شاخهٔ کمینه‌گرای نمونه‌سازیهای شناختی زبان در نظر گرفت، این پایانه‌ها سازشی را نشان می‌دهند که روشهای شناختی در جهت تلاش برای نمونه‌سازی مغز انسان پذیرفته‌اند. با وجود این، پایانه‌های «زبان طبیعی» به لحاظ حیطهٔ زبانی که پوشش می‌دهند به هیچ وجه رضایت‌بخش نیستند.

از سوی دیگر، دیدگاه کمینه‌گرا می‌کوشد تا پدیده‌های مجزای مغز را نمونه‌سازی کند؛ سپس با کنار هم قرار دادن آن پدیده‌ها به کل آن نظام، یعنی به اهداف آن یا صورت مشابه سازی شدهٔ آن دست یابد. یکی از نمونه‌های این گرایش نسبت به پدیدهٔ مغز، روشهای پردازش زبان بر پایهٔ احتمالات[۳۲] است. در این نوع پردازش، تدابیر شناختی مغز به هیچ وجه مورد توجه نیست، بلکه سعی می‌شود پدیدهٔ زبان را با توصیفی آماری توضیح داد. در این روش، پدیده‌ها نمونه‌سازی می‌شود، نه نظامها و آنچه اهمیت دارد رایانه‌ای کردن کارکردهای این پدیده‌هاست؛ پدیده‌هایی که برخلاف دیدگاه کل‌گرا به وسیلهٔ رابط مشترکی به نظام مورد مطالعه یعنی مغز پیوند نخورده است. در این راستا، حتی ممکن است جریاناتی رخ دهد که اساساً متفاوت از فرایندهای درونی مغز باشد.

مزایا و معایب گرایش کمینه‌گرا نسبت به پدیده‌های مرتبط با مغز کاملاً برعکس مزایا و معایب مطالعات کل‌گراست. یکی از مزایای نظریهٔ زبانی مبتنی بر احتمالات یا به عبارتی کمینه‌گرا آن است که این نظام با فرض در اختیار داشتن پیکره‌ای کافی از داده‌های

نمونه‌گیری شده[۳۳] (داده‌های نمونه)، با سرعتی بیش از نظامهای شناختی مشابه خود به نتیجه می‌رسد. با وجود این، در گرایشهایی که نسبت به پدیده‌های مغزی بر مبنای احتمالات عمل می‌کنند، سهم عمده‌ای از تلاشها تکراری است؛ زیرا این گرایشها، آن پدیده‌ها را ناشی از نظامی واحد یعنی مغز نمی‌دانند، نظامی که داده‌های تمامی این رشته‌ها را می‌توان در آن جستجو کرد. بنابراین، نه تنها بررسی زبان از دیدگاه احتمالات، حداقل به لحاظ نظری به تلاشهای تکراری می‌انجامد، بلکه کنش متقابل میان رشته‌هایی که این پدیده‌ها را مطالعه می‌کنند به حداقل می‌رسند.

دیدگاه کل‌گرا نسبت به زبان، ذهن را مانند ماشینی در نظر می‌گیرد که پدیده‌های گوناگون از آن سرچشمه می‌گیرد. این دیدگاه نسبت به زبان، ذهن را پدیده مجزایی تلقی می‌کند که داده‌های حوزه‌های گوناگون مطالعاتی را فراهم می‌کند؛ حوزه‌هایی که میان آنها ارتباطی وجود ندارد. برای پژوهشگر کمینه‌گرا ذهن مترادف است با فرایند مورد مطالعه‌اش؛ یعنی برای او ذهن فرایند است.

دسته‌بندی زبانهای برنامه‌نویسی

زبان برنامه نویسی زبانی ساختگی[۳۴] است که برای کنترل رفتار ماشین، به خصوص رایانه‌ها، به کار می‌آید. مانند زبانهای طبیعی با به کارگیری قواعد نحوی و معنایی که به ترتیب به منظور تعیین ساختار و معنا به کار می‌آیند، قابل تعریف‌اند. زبانهای برنامه‌نویسی بسیار فراوان و متنوع بوده و تعدادشان پیوسته رو به افزایش است. درمیان زبانهای برنامه‌نویسی نوعی تمایز میان زبانهای رویه‌ای[۳۵] (برای مثال، سی[۳۶] و پاسکال[۳۷] و زبانهای اعلانی[۳۸] (برای مثال، پرولوگ[۳۹]) مطرح است. زبانهای رایانه‌ای اعلانی آن دسته از زبانهایی‌اند که در واقع مشابه فرایندهای استدلالی مغز انسان عمل می‌کنند. در زبان‌شناسی رایانه‌ای معمولاً روندهای شناختی با زبانهای رایانه‌ای[۴۰] اعلانی کدنویسی می‌شود، در حالی که مطالعات آماری نسبت به کاربرد زبانهای برنامه‌نویسی رویه‌ای گرایش دارند؛ زیرا این زبانها به صورت استاندارد شامل توابع ریاضی‌اند و با برنامه نویسی محاسبات آماری همسوترند. این تمایز را حتی می‌توان در مطالعهٔ زبان دید. برای مثال، دیدگاه اعلانی نسبت به زبان آن است که هر عبارتی می‌تواند شامل، برای مثال پنج عامل فاعل، مفعول، متمم، قید و فعل باشد. در مقابل، دیدگاه رویه‌ای نسبت به زبان باید علاوه بر آن بداند که کدام عامل نخست پردازش می‌شود (اویی،[۴۱] ۱۹۹۸:۱۸). حال زبانهای برنامه‌نویسی اعلانی[۴۲] و رویه‌ای کدام‌اند و چه تفاوتهایی با یکدیگر دارند؟

برنامه نویسی اعلانی

در برنامه‌های اعلانی به معنای فرمانها توجه می‌شود، نه به توالی[۴۳] آنها. در این نوع برنامه‌ها می‌توان کدنویسی را با هر ترتیبی از توالی فرمانها انجام داد. بنابراین، هدف برنامه‌نویسان برنامه‌های اعلانی دستیابی به پاسخ است، نه سلسله مراتب رسیدن به آنها. در این برنامه مهم نیست که پاسخ با چه توالی خاصی از فرمانها به دست آمده است؛ آنها برای دستیابی به هدف، داده‌ها را گرفته، سپس زنجیره‌ای از راه کارهای گوناگون را یک به یک می‌آزمایند. هنگامی که یک راه کار رد می‌شود، «پس رفت»[۴۴] رخ می‌دهد و بدین ترتیب برنامه به مرحلهٔ وقوع آخرین تصمیم‌گیری باز می‌گردد و از آنجا مسیر جدیدی انتخاب می‌شود که هنوز آزموده نشده است. اگر هیچ یک از مسیرهای ممکن پذیرفته نشود، برنامه نادرست است. بنابراین، برنامه‌های اعلانی متکی بر ترتیب رویّه‌های[۴۵] خود نیستند، زیرا با هر ترتیبی که در برنامه قرار گیرند به هر حال تا دستیابی به هدف، یک به یک آزموده خواهند شد.

برنامه‌های اعلانی تا حدودی عملکردهای ذهن انسان را الگوبرداری می‌کنند. جریانهای استدلالی ذهن ما بیشتر اعلانی‌اند؛ یعنی در آنها توالی اهمیت چندانی ندارد. به عبارت دیگر، هنگامی که فردی به پرسشی پاسخ می‌دهد، به توالی سازه‌هایی که برای ارائه آن پاسخ در ذهنش طی شده است اهمیت نمی‌دهد. برنامه‌نویسی اعلانی هدف‌گرا[۴۶] است و در آن از داده‌ها به کرات استفاده می‌شود؛ زیرا می‌توان تغییرات ایجاد شده در داده‌ها را وارونه کرد. بنابراین، در برنامه‌نویسی اعلانی پرداختن به داده‌های جامع مشکل آفرین نیست، زیرا می‌توان از آنها هر چند بار که لازم باشد استفاده کرد؛ یعنی تا زمان دستیابی به هدف یا آزمودن تمام مسیرهای دستیابی به آن و پی بردن به نادرستی همهٔ آنها.

اما آیا زبان برنامه نویسی اعلانی، آن گونه که تعریف شد، وجود دارد؟ باید در صحت این مطلب تردید داشت. به صراحت می‌توان گفت که برخی از زبانهای برنامه‌نویسی هوش مصنوعی، برای مثال، لیسپ[۴۷]، که جامعهٔ علوم رایانه‌ای آنها را زبان اعلانی می‌داند در واقع چنین نیستند. گرچه عبارت «Defun» در زبان برنامه‌نویسی لیسپ سازه‌های اعلانی دارد، این زبان تا حدود زیادی رویّه‌ای است. تصور عمومی بر آن است که لیسپ یک زبان برنامه نویسی اعلانی است و احتمالاً همین تصور منشأ بسیاری از برداشتهای نادرست در مورد این زبان است. زبان برنامه‌نویسی پرولوگ تنها زبان پرکاربردی است که به معنای واقعی دارای توانش اعلانی است؛ با وجود این، می‌توان حتی پرولوگ را نیز با استفاده از سازه‌های نیرومندی که در آن تعبیه شده است به صورت رویّه‌ای نوشت. همچنین می‌توان در برنامه پرولوگ از شرط «else» استفاده کرد. برای مثال،

می‌توان گزاره‌ای نوشت که بدون هیچ شرطی برای آزمودن، به طور مستقیم، فرمان را انجام دهد؛ اما در آن صورت دیگر برنامه کاملا اعلانی نیست، زیرا این گزینه که هیچ گونه شرطی برای آزمودن آن نیست مانع از حرکت برنامه در میان سایر گزینه‌ها می‌شود. بنابراین، هرچند پرولوگ یک زبان برنامه نویسی کاملاً اعلانی است، با به کارگیری تدابیر مؤثر کدنویسی، می‌توان در این زبان نیز کدهایی نوشت که ماهیت رویّه‌ای داشته باشند.

به اجمال می‌توان گفت، زبانهای برنامه نویسی اعلانی هدف‌گرا هستند؛ در آنها توالی انجام دادن فرمانهایی که منتهی به دستیابی به هدف می‌شود اهمیت ندارد و آنچه مهم است انجام دادن این هدف با هر یک از مسیرهای ممکن است. باید اذعان داشت همان طور که برخی از زبانهای برنامه‌نویسی اعلانی‌اند، اغلب به صورت رویّه‌ای به کار می‌روند. برای نمونه می‌توان به این نوع زبانهای برنامه‌نویسی اعلانی اشاره کرد: Poplog، SQL، Metapost، Analytica، ABSET و Visual Prolog.

برنامه نویسی اعلانی فرایندهای شناختی

در پژوهشهای کل‌گرای هوش مصنوعی به کارگیری برنامه نویسی اعلانی، مزیت مهمی است. تأکید نداشتن بر توالی فرمانها دو فایده دارد: نخست آنکه تاحدودی امکان کارکردن دو یا چند برنامه نویس را روی برنامه فراهم می‌کند، زیرا اگر خطایی در برنامه نویسی رخ دهد، تغییرات شدید و ناگهانی در معنای برنامه رخ نمی‌دهد؛ البته با این فرض که برنامه سراسر اعلانی باشد. دوم آن که، برنامه نویس فقط باید فرایندهای شناختی‌ای را که مشخص شده است کدنویسی کند و نیازی به آگاهی از توالی فرایندها در ذهن انسان ندارد. این مورد، مزیت مهمی است، زیرا مطالعات شناختی در آن حد پیشرفت نکرده که پژوهشگران این رشته بتوانند راجع به توالی کنشهای مغز استدلال کنند. بنابراین، گرایشی که به چنین اطلاعاتی نیاز دارد، اکنون عملی نیست. در نتیجه، برنامه‌نویسی اعلانی عمیقاً با دیدگاه کل‌گرا نسبت به فرایند زبان سازگار است؛ زیرا مشخصه‌های برنامه نویسی اعلانی به خوبی بر مشخصه‌های مورد نیاز برای پیروی از پاسخ کل‌گرا با پدیدهٔ زبان منطبق است.

برنامه نویسی رویه‌ای

زبانهای رایانه‌ای رویه‌ای که در حقیقت بیشتر زبانهای رایانه‌ای را تشکیل می‌دهند، به توالی فرمانها می‌پردازند تا به دستیابی صرف به هدف نهایی. توالی فرمانها بر معنای برنامه‌های رویّه‌ای تأثیر بسیاری دارد. در این برنامه‌ها سعی می‌شود با استفاده از داده‌ها، مسیر حرکت برنامه محدود شود، یعنی برعکس برنامه‌های اعلانی که در آنها از داده‌های فراهم شده

برای یافتن مسیرهای موجود استفاده می‌شود. در این برنامه‌ها، معمولاً پس رفت وجود ندارد، اما از تمامی داده‌های موجود برای محدود کردن گزینه‌های ممکن استفاده می‌شود. فرمانها فقط هنگامی انجام می‌شود که تمامی شرطها پذیرفته شود. در برنامه‌های رویه‌ای، کاربرد جملات شرطی «else» بسیار معمول است، درست به همان اندازه که توالی فرمانها اهمیت ضروری دارد. اگر شرط «else» در محل نادرستی قرار بگیرد، معنای برنامه تغییر می‌کند. بنابراین، کدنویسی اطلاعات به صورت رویه‌ای مستلزم آگاهی از توالی دقیق کنشهاست. با این توالی، مجموعه مسیرهای مشخصی در میان برنامه مرزبندی می‌شود و جهت حرکت برنامه در این مسیرها تعیین می‌گردد.

برنامه‌های رویه‌ای فرمانهای متوالی را تا پایان تعقیب می‌کنند. این برنامه‌ها تا رسیدن به هدفی که به صورت اعلانی تعریف شده باشد متوقف نمی‌شوند. در این برنامه‌ها باید محدودیتهای[۴۸] مناسبی را گنجاند، محدودیتهایی که فعال‌سازی کدهای داده‌های ورودی و غیر فعال سازی کدهای نامربوط را تضمین کند. بنابراین، برنامهٔ اعلانی پس از انجام هدفی که به صورت اعلانی تعریف شده باشد خاتمه می‌یابد؛ حال آنکه برنامه رویه‌ای برای دستیابی به هدف، در میان خطوط فرمان به صورت متوالی حرکت می‌کند و تا تکمیل آخرین خط فرمان هیچ پس رفتی انجام نمی‌دهد. زبانهای برنامه نویسی Cold Fusion ، Basic ، C++ ، C ، Visual Basic ، Perl ، Java ، Fortran ، Java Script ، Script ، Action ، Delphi ، Pascal و بسیاری از زبانهای برنامه‌نویسی دیگر، همگی نمونه‌هایی از زبانهای برنامه‌نویسی رویه‌ای‌اند.

برنامه‌نویسی رویه‌ای در پدیده زبان

از نظر عملی، روشهای مبتنی بر احتمالات (روشهای احتمالاتی)[۴۹] بیشتر با زبانهای برنامه‌نویسی رویه‌ای سازگارند؛ زیرا این زبانها بهتر از زبانهای اعلانی می‌توانند به توابع ریاضی بپردازند. البته این امر نه به دلیل ناتوانی اصولی زبانهای برنامه‌نویسی اعلانی، بلکه بیشتر چون رشته‌هایی که زبانهای رویه‌ای را طی سالها گسترش داده‌اند بیشتر مبنای ریاضی داشته‌اند؛ در نتیجه زبانهایی مانند سی، پاسکال، وی بی[۵۰] و بسیاری از زبانهای دیگر، به صورت استاندارد شامل توابع ریاضی‌اند. اما در زبانی مانند پرولوگ باید چنین توابعی را از مرحلهٔ پیش‌نویس تهیه کرد. بنابراین، در دیدگاهی که نسبت به فرایند زبان کمینه‌گرای احتمالاتی است، به کارگیری زبان برنامه نویسی سازگار با اصول ریاضیات دارای مزیتهای فراوانی است. از آنجا که برنامه نویسی اعلانی بیشتر بر پایه اصول شناختی عمل می‌کند تا ریاضیاتی، در نتیجه برای کدنویسی اطلاعات در روشهای احتمالاتی باید از میان زبانهای رویه‌ای که پایهٔ ریاضی‌اند، یکی را برگزید.

به لحاظ نظری، روشهای احتمالاتی در جستجوی یافتن توالیهای وقوع در متن/گفتارند. این نوع توالی برخلاف توالی پدیده‌های درون ذهنی که مورد توجه شناخت گرایان است، به طور مستقیم مشاهده و بررسی می‌شود. انگاره مارکو[۵۱] از جمله روشهای احتمالاتی است که با آن می‌توان توالیهای موجود در متن یا گفتار را بررسی کرد. داده‌هایی که از فرایند مارکو[۵۲] عبور می‌کنند، به لحاظ توالی اجزاء، تجزیه و تحلیل می‌شوند. این روش تحلیل داده- راندنی،[۵۳] به طور مستقیم بر روند درک و تولید زبان اثر می‌گذارد. از آنجا که روش‌شناسی مقدماتی با گرایش کمینه‌گرای احتمالاتی ماهیت رویه‌ای دارد، برای برنامه نویسی در این گرایش باید زبانهایی را برگزید که به لحاظ نظری به این دیدگاه نزدیک‌تر باشند؛ یعنی یکی از زبانهای رویه‌ای.

زبان‌شناسی رایانه‌ای: رشتهٔ دو شاخه‌ای

به طور کلی، زبان‌شناسی رایانه‌ای به دو بخش تقسیم می‌شود: بخش نخست که زبان‌شناسی رایانه‌ای را فلسفهٔ کاربردی در نظر می‌گیرد، نسبت به فرایند زبان وابسته به نگرش کل‌گرا است. از این نظر، نظریات موجود، خواه توصیفی یا غیر توصیفی، باید برای اشاره به کل زبان در چهارچوب یک نظریهٔ زبانی عمل کنند. نگرش کل‌گرا نسبت به فرایند زبان، خود مستلزم پیروی از گرایش شناختی است؛ زیرا نگرش کل‌گرا تنها توجه به نظامی که پدیده مورد مطالعه را ایجاد کرده است شکل می‌گیرد. این امر خود مستلزم توجه به روندهای استدلال کنشهای مغز و نیز گرایش به رشته‌ای است که آن را مطالعه می‌کند، یعنی همان «علم» شناخت. تمایل به فرایند شناختی، منجر به رویکردی می‌شود که متأثر از زبانهای اعلانی است؛ زیرا رویکرد رویه‌ای به فرایندهای شناختی برای محدود سازی مؤثر عملکرد برنامه داده‌های کافی در اختیار ندارد و این ناشی از این است که توالی فرایندها در ذهن انسان دیدنی نیست. نگرش کل‌گرا فوایدی دارد؛ از جمله: پژوهشگرانی که پدیده‌های متفاوت را از دیدگاه کل‌گرایی مطالعه می‌کنند، می‌توانند به واقعیاتی در مورد فرایندهای شناختی و نظامهایی اشاره کنند که پژوهشگر کل‌گرای زبان طبیعی آن را بررسی می‌کند. مهم‌ترین کاستی این گرایش نیاز به کدنویسی داده‌های فراوان است. اگرچه این مشکل با به کارگیری فرهنگهای ماشین خواندنی[۵۴] و شیء‌گرایی[۵۵] تا حدودی برطرف می‌شود، در مسیر برنامه‌نویسی اعلانی همچنان باقی می‌ماند.

بخش دوم از زبان‌شناسی رایانه‌ای فقط به آنچه از طریق رایانه صورت می‌گیرد، محدود می‌شود و این خود مستلزم گرایش کمینه‌گراست؛ زیرا اعمال گرایش کل‌گرا به زمانی

طولانی نیاز دارد. نگرش کمینه‌گرا نسبت به زبان عمدتاً با پردازش زبان بر پایه احتمالات (پردازش احتمالاتی زبان)[۵۶] اعمال می‌شود. پردازش احتمالاتی زبان، امکان تبادل اطلاعات میان حوزه‌هایی را که پدیده‌های مرتبط با مغز را مطالعه می‌کنند، در حد تبادل اصطلاحات روشهای احتمالاتی محدود می‌کند. این امر بدان سبب است که بین حوزه‌های مطالعاتی، بخش مشترکی وجود ندارد و پدیده‌های مورد مطالعه به صورت مجزا و بدون در نظر گرفتن نظام پدید آورنده بررسی می‌شوند. علاوه بر این، پردازش احتمالاتی زبان مستلزم کاربرد زبانهای برنامه‌نویسی رویه‌ای است. این زبانها نه تنها از پیش نیازهای ریاضیاتی گرایشهای احتمالاتی پشتیبانی می‌کنند، بلکه به لحاظ نظری به الگوهای رویه‌ای مبتنی بر احتمالات که برای پردازش زبان در این گرایش از آن استفاده می‌شود شباهت دارند. نخستین مزیت گرایشهای احتمالاتی آن است که داده‌های مورد نیازشان واحدهایی خلاصه شده و بسیار دقیق هستند؛ بنابراین تعیین ویژگیهای آنها ساده‌تر است. این سادگی بدان سبب است که هر «دانش» یا «درکی» که به واژه مربوط می‌شود، طبق روشهای آماری دقیق محاسبه می‌گردد و نیاز نیست که با ارجاع به یک نظریهٔ شناختی کدگذاری شود. دومین مزیت گرایشهای احتمالاتی آن است که در آنها، روشهای پرداختن به داده‌ها، الگوهای صوری بسیار پیشرفته‌ای دارند؛ بنابراین، می‌توان برای نتایج آن دلایل صوری ارائه کرد. گزدر[۵۷] نیز در کتاب *کاربردشناسی* (گزدر، ۱۹۷۹) به این مطلب اشاره کرده است. عمده‌ترین کاستی گرایش احتمالاتی آن است که اطلاعاتی که پس از بررسی هر یک از پدیده‌های مغزی تولید و کدنویسی می‌شوند، هیچ گونه کاربردی برای پژوهشگرانی که سایر پدیده‌های مغزی را مطالعه می‌کنند ندارد. به دلیل ماهیت کمینه‌گرای این آثار، پیوندی میان بررسی یک پدیده و پدیدهٔ دیگر وجود ندارد. پیامد این امر آن است که نگرش کمینه‌گرا، حداقل به لحاظ نظری، با دوباره کاری مواجه است.

نتیجه‌گیری

به طور کلی، با آغاز دهه ۶۰، هنگامی که نظریات چامسکی به تدریج دیدگاه غالب در زبان‌شناسی محسوب شد، بررسی توصیفی زبان به فراموشی سپرده شد. در این دوران زبان شناسان نظریه پرداز(زبان‌شناسان پشت میز نشین) چون فیل‌مور، روش‌شناسی پیکره را کاملا بی اعتبار خواندند (میر،[۵۸] ۴:۲۰۰۲). از آنجا که چامسکی نمونه شاخص دیدگاه شناختی است، تعجب‌آور نیست اگر پس از دهه ۶۰ زبانهای برنامه‌نویسی اعلانی، به ویژه زبان پرولوگ، به عنوان زبانهای برنامه‌نویسی غالب در زبان‌شناسی رایانه‌ای مطرح شوند. این

مطلب را می‌توان با مشاهده منابع بسیار متنوعی که پس از دهه ۶۰ در زبان‌شناسی رایانه‌ای، با استفاده از زبان برنامه نویسی پرولوگ نوشته شده است به سادگی دریافت. اما در اواخر دهه ۷۰ به تدریج مشخص شد که نظریهٔ چامسکی گنجینهٔ اندیشمندانه‌ای که در آغاز چنین می‌نمود نیست و زبانهای برنامه‌نویسی رویه‌ای، گرایش ریاضی را در زبان‌شناسی ترجیح می‌دادند، دوباره دیدگاههای توصیفی قوت گرفت و روش‌شناسی پیکره که پایه توصیف تجربی زبان است گسترش یافت. البته پیشرفتهای علوم رایانه‌ای، یعنی افزایش سرعت و قدرت رایانه‌ها، این امکان را فراهم ساخت که روشهای توصیفی به نتایج خیره کننده‌ای دست یابد. همچنین، امکان ذخیره‌سازی پیکره‌های بسیار بزرگ و انواع جستجو در میان داده‌های آن فراهم آمد. این موضوع خود باعث گسترش روشهای پردازش زبان بر مبنای اصول آمار و احتمالات بود. بدین ترتیب با آغاز دهه ۸۰ به تدریج زبانهای برنامه‌نویسی رویه‌ای، به خصوص زبان برنامه‌نویسی «سی» در پژوهشهای زبان‌شناسی رایانه‌ای رواج یافت.

بنا بر آنچه آمد، می‌توان گفت که انتخاب و آموزش زبان برنامه نویسی مناسب با پژوهش مورد نظر از اهمیت بسزایی برخوردار است. در یک جمع بندی کلی و با توجه به آنچه اکنون در زمینه پردازش زبان در دنیا مطرح است - منظور گسترش پژوهشهای پیکره بنیان، ضرورت تحلیلهای آمار و احتمالات و سایر تحلیلهای ریاضی بنیان در زبان‌شناسی رایانه‌ای و مسائلی از این دست - شاید بتوان گفت برای پژوهشگران زبان‌شناسی رایانه‌ای، آشنایی با زبانهای برنامه‌نویسی رویه‌ای کاربردی‌تر است. اما به هر حال، دانستن زبان برنامه نویسی‌ای مانند پرولوگ هم می‌تواند علاوه بر کاربرد در گرایشهای شناختی، در زمینه درک منابع غنی دهه‌های گذشته این شاخه میان رشته‌ای مفید واقع شود. گفتنی است زبانهای برنامه نویسی که در واقع روش تفکر رایانه‌ها را شرح می‌دهند، از اصول مقدماتی مشترکی پیروی می‌کنند. این امر باعث می‌شود که با دانستن یک زبان برنامه نویسی، بتوان میزان زیادی از فرمانهای سایر برنامه‌ها را درک کرد.

نکته آخر آن که بهره‌گیری مؤثر از منابع درسی زبان‌شناسی رایانه‌ای نیازمند معادل سازی اصطلاحات تخصصی این رشته در زبان فارسی، گزینش و معرفی منابع سودمند، گردآوری و ترجمه این منابع در زبان فارسی، آشنایی دانش آموختگان زبان‌شناسی با علوم رایانه‌ای، به خصوص مبانی و اصطلاحات رایانه و برنامه‌نویسی، است. برای رسیدن به این هدف باید نیروی انسانی متخصص در این رشته تربیت شود و ساعات آموزشی این درس افزایش یابد. با توجه به آنکه زبان‌شناسی رایانه‌ای موضوعی میان رشته‌ای و پرکاربرد است، لذا توجه مراکز آموزشی مربوط به تربیت نیروی انسانی متخصص در این شاخه می‌تواند بسیار سودمند باشد.

پی‌نوشتها

1. computational linguist
2. Essex, Edinburgh and Washington.

۳. برای نمونه می‌توان به دانشگاه استکهلم (Stockholm) اشاره کرد.

۴. حتی در برخی از کشورهای خاور دور، موضوع پردازش زبان از مقطع دبیرستان تدریس می‌شود که البته به نظرمی رسد این مسئله بیشتر به دلیل موانع فراوانی باشد که پردازش زبانهای خاور دور با آن مواجه است.

۵. دانشجویان زبان شناسی از رشته‌های تحصیلی متنوعی وارد این رشته می‌شوند. بنابراین گذراندن پیش‌نیازهای خاصی برای همگان الزامی می‌گردد.

6. abbreviation
7. commands

۸. برای مثال، در گروه زبان‌شناسی دانشگاه استکهلم، ۸ واحد آموزشی به تدریس برنامه نویسی در دو مرحله می‌پردازد: مقدماتی و پیشرفته اختصاصی.

9. Hutchins
10. finite state programming

۱۱. در این نوع برنامه‌نویسی، زبان شناس واقعیات مربوط به زبانی را که در حال الگوسازی آن است کدنویسی می‌کند.

12. Beesley
13. Mc Enery

۱۴. گفتنی است که در این بحث مکنری بیشتر بر نظامها و آنچه نمونه‌سازی می‌شود تأکید می‌کند؛ ضمن آنکه به گونه‌ای جانبی به مسئلهٔ انتخاب زبان برنامه‌نویسی مناسب می‌پردازد. اما در ترجمه‌ای که مطالعه می‌کنید با استفاده از عنوانها و افزودن بخشهای مقدمه و نتیجه‌گیری مسئله بر انتخاب زبان برنامه‌نویسی مناسب بیشتر تأکید شده است.

15. Mitcov
16. Jurafsky
17. natural language processing
18. speech recognition
19. computational psycholinguistics
20. artificial intelligence
21. cognitive approach
22. probabilistic approach
23. holistic
24. reductionist
25. intelligent systems
26. Chomsky
27. Shannon and Weaver
28. holisticism
29. reductionism
30. natural language interface

31. natural language front ends
32. probabilistic language processing techniques
33. sample data
34. artificial language
35. procedural languages
36. C
37. PASCAL
38. declarative languages
39. PROLOG(PROgramming in LOGic)
40. computer languages
41. Ooi
42. declarative programming
43. sequence
44. backtracking

۴۵. procedure : برنامه یا زیر برنامه‌ای است که خود به عنوان جزئی از یک برنامه بزرگ‌تر به کار می‌رود یا می‌تواند به کار رود.

46. goal driven
47. LISP
48. constraints
49. probabilistic techniques
50 VB(Visual Basic)
51. Markov model
52. Markov process
53. data driven
54. machine readable dictionaries
55. object orientation
56. probabilistic natural language processing
57. Gazder
58. Meyer

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

منابع

Beesley, K. and Karttunen, L. (2003), *Finite State Morphology*. CSLI Publications, United States.

Chomsky, N.(1965), *Aspects of the Theory of Syntax*, MIT Press, Cambridge (Mass). as Cited in Mc Enery, T.(1992), *Computational Linguistics*, SIGMA Press, Cheshire, England.

Gazder, G. (1979), *Pragmatics, Implicature, Presupposition and Logical Form*, Academic Press, New York, as cited in Mc Enery, T.(1979), *Computational Linguistics*, SIGMA press, Cheshire, England, 1992.

Hutchins, W.J.(1986), *Machine Translation: Past, Present, Future*. Ellis Horwood Limited, Great Britain.

accessed on: http://ourworld.compuserve.com/homepage/WJHutchines/PPF-19.pdf 2007-02-05

Jurafsky, D. and J.H. Martin (2000), *Speech and Language Processing*, Prentice-Hall, New Jersey.

Mc Enery, T., (1992), *Computational Linguistics*, SIGMA Press, Cheshire: England.

Meyer, C. F.. (2002), *English Corpus Linguistics*, Cambridge: Cambridge University Press.

Mitkov, R. (2003), *The Oxford Handbook of Computational Linguistics*, New York: Oxford University Press.

Ooi, V. B. Y. (1998), *Computer Corpus Lexicography*, Edinburgh: Edinburgh University Press.

Shannon, C.E. and W. Weaver (1949), *The Mathematical Theory of Communication*, Urbana: University of Illinois Press, as cited in Mc Enery, T. (1992), *Computational Linguistics*, Cheshire, England: SIGMA press.

دبیر مقدم، محمد (۱۳۷۸)، *زبان‌شناسی نظری پیدایش و تکوین دستور زایشی*، تهران: انتشارات سمت.

عاصی، مصطفی و محمد عبدعلی (۱۳۷۵)، *واژگان گزیده زبان‌شناسی*، تهران: شرکت انتشارات علمی فرهنگی.

محمدی‌فر، محمدرضا (۱۳۸۳)، *واژه نامهٔ کامپیوتر*، ویراست دوم، تهران: فرهنگ معاصر.

همایون، همادخت (۱۳۷۹)، *واژه نامه زبان‌شناسی و علوم وابسته*، تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی