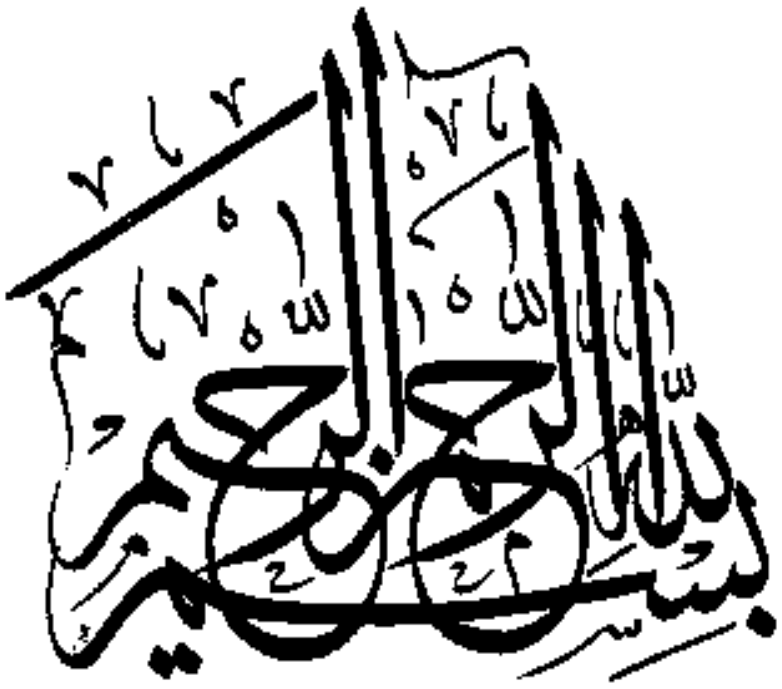

بسم الله الرحمن الرحيم

دانشگاه صنعتی امیرکبیر
دانشکده مهندسی پزشکی

پایان نامه کارشناسی ارشد

(گرایش بیوالکتریک)

عنوان:

مدلسازی عصبی عضلانی سیستم تولیدگفتار

استاد راهنما: دکتر فرزاد توحیدخواه

استاد مشاور: دکتر شهریار غریب‌زاده

دانشجو: ایوب دلیری

زمستان ۸۶

بسمه تعالی

شماره:

تاریخ :

معاونت پژوهشی
فرم پروژه تحصیلات تکمیلی ۷

فرم اطلاعات پایان نامه
کارشناسی ارشد و دکترا

۱-مشخصات دانشجو

نام و نام خانوادگی: ایوب دلیری دانشجوی ارشد ☑ بورسیه ☐ معادل ☐

شماره دانشجویی ۸۴۱۳۳۰۶۷: دانشکده : مهندسی پزشکی رشته تحصیلی : مهندسی پزشکی-بیوالکتریک

نام و نام خانوادگی استاد راهنما : فرزاد توحیدخواه

عنوان به فارسی : مدلسازی سیستم عصبی-عضلانی تولید گفتار

عنوان به انگلیسی : Speech motor control modeling

نوع پروژه: کاربردی ☐ بنیادی ☐ توسعه ای ☑ نظری ☐

تاریخ شروع : ۸۵/۹/۱۴ تاریخ خاتمه : ۸۶/۱۰/۱۵ تعداد واحد : ۹

سازمان تأمین کننده اعتبار :

واژه های کلید به فارسی :
تولید گفتار، عصبی عضلانی، کنترل پیش‌بین مبتنی بر مدل، مخچه، بروکا.

واژه های کلیدی به انگلیسی :
Speech production, model control predictive, cerebellum, broca's area.

نظرها و پیشنهادها به منظور بهبود فعالیت های پژوهشی دانشگاه :

استاد راهنما :

دانشجو :

امضاء استاد راهنما : تاریخ : ۸۷/۲/۱۴

نسخه ۱ : معاونت پژوهشی
نسخه ۲ : کتابخانه و به انضام دو جلد پایان نامه به منظور تسویه حساب با کتابخانه و مرکز اسناد و مدارک علمی

چکیده

از آنجا که گفتار مهمترین راه ارتباطی برای انسان به شمار می‌آید، بسیار مورد مطالعه قرار گرفته شده است. برای تولید گفتار مدل‌های مختلفی ارایه گردیده است که می‌توان آن‌ها را به دو دسته عمده تقسیم نمود. دسته اول عموما دارای پایه و اساس فیزیولوژیکی نیستند و صرفا با استفاده از سیگنال گفتار و روش‌های پردازشی، مدل‌هایی برای تولید گفتار ارایه می‌دهند. دسته دوم به مبحث زبان و تولید و درک گفتار در سطوح عالی مغز می‌پردازند چرا که گفتار مهمترین راه خروجی زبان است. مدل‌هایی که در دسته دوم ارایه شده‌اند عموما کیفی هستند و بیشتر برای بررسی و پیش‌بینی تاثیر آسیب‌های مغزی بر روی زبان و گفتار بکار می‌روند. نکته‌ای که کمتر به آن توجه شده است این است که تولید گفتار در واقع مجموعه‌ای از اعمال عصبی عضلانی است. بر طبق مشاهدات و آزمایشات، سیستم کنترل تولید گفتار در سطوح مختلفی به پیش‌بینی اغتشاشات می‌پردازد. همچنین شواهد حاکی از وجود مدل‌های داخلی از سیستم حرکتی تولید گفتار در مغز است و سیستم عصبی بر اساس این مدل‌ها به کنترل آرتیکولاتورهای گفتار می‌پردازد. بر اساس این مشاهدات در این پروژه برای سیستم کنترل تولید گفتار مدلی پیشنهاد شده‌است که بر اساس کنترل پیش‌بین مبتنی بر مدل است. شبیه‌سازی‌ها نشان دادند که مدل پیشنهادی می‌تواند داده‌های حرکتی بدست آمده از سیگنال گفتار را به خوبی تولید کند و این مدل مشابه سیستم تولید گفتار به خوبی اغتشاشات خارجی را حذف می‌کند. همچنین بازسازی سیگنال گفتار ناشی از حرکات کنترل شده به وسیله مدل پیشنهادی نشان داد که این مدل می‌تواند رفتار سیستم تولید گفتار را به خوبی توصیف کند.

کلید واژگان:

تولید گفتار، عصبی عضلانی، کنترل پیش‌بین مبتنی بر مدل، مخچه، بروکا.

فهرست مطالب:

۱-مقدمه ۲

۱-۱- مدلسازی تولید گفتار ۳

۱-۲- مراکز مغزی مرتبط با تولید گفتار ۴

۱-۳- شواهد وجود پیش‌بینی در سیستم تولید گفتار ۴

۱-۴- مدل پیشنهادی برای این پروژه ۶

۱-۵- اهداف پروژه و روش انجام آن ۷

۱-۵-۱- اهداف پروژه ۷

۱-۵-۲- فرضیات مدلسازی ۸

۱-۵-۳- استخراج داده‌ها ۸

۱-۵-۴- پیاده سازی کامپیوتری ۸

۱-۶- روند تنظیم پایان نامه ۹

۲- فیزیولوژی سیستم تولید گفتار ۱۱

۲-۱- مراکز مغزی مرتبط با گفتار ۱۱

۲-۱-۱- نواحی قشری مرتبط با گفتار ۱۱

۲-۱-۲- سازماندهی نواحی ارتباطی پیکری، شنوایی و بینایی ۱۴

۲-۱-۳- عملکرد مغز در بر قراری ارتباط زبانی ۱۶

۲-۲-اندام‌های مسیر صوتی ۱۹

۲-۲-۱- شش‌ها و نای ۲۱

۲-۲-۲- حنجره و تارهای صوتی ۲۲

۲-۲-۳- فضای دهان وحلق ۲۷

۲-۳- خلاصه مطالب ۲۹

۳- مدلسازی سیستم تولید گفتار ۳۱

۳-۱- مدلسازی عملگرهای سیستم تولید ۳۲

۳-۱-۱- انتشار صدا و مدلسازی آن ۳۴

۳-۱-۲- انواع تحریک مجرای صوتی ۳۷

۳-۱-۳- تحریک واکدار ۳۸

۳-۱-۴- منابع تحریک بی‌واک‌ها ۴۱

۳-۲- مدلسازی سیستم تولید گفتار از دیدگاه عصبی ۴۱

۳-۲-۱- مدل ورنیکه-بروکا ۴۲

۳-۲-۲- مدل‌های جدید برای زبان ۴۴

۳-۲-۳- مدلی محاسباتی برای تولید گفتار ۴۵

۳-۳- خلاصه مطالب ۴۷
۴- مدل پیشنهادی برای سیستم عصبی عضلانی تولید گفتار ۴۹
۴-۱- معرفی کنترل پیش بین مبتنی بر مدل ۴۹
۴-۱-۱- استراتژی کنترل پیش‌بین ۵۰
۴-۲- مدل پیشنهادی برای سیستم کنترلی تولید گفتار ۵۲
۴-۲-۱- شواهد وجود پیش‌بینی در سیستم کنترلی تولید گفتار ۵۲
۴-۲-۲- تناظرات کنترل پیش‌بین و سیستم کنترلی تولید گفتار ۵۴
۴-۳- مدل بیومکانیکی و سنتزکننده گفتار ۵۹
۴-۳-۱- لوله آکوستیکی بدون اتلاف ۵۹
۴-۳-۲- مدل منبع تحریک ۶۱
۴-۳-۳- تکمیل مدل لوله آکوستیکی ۶۳
۴-۴- مدل پیشنهادی در این پروژه ۶۵
۴-۵- خلاصه مطالب ۶۷
۵- شبیه‌سازی مدل پیشنهادی ۶۹
۵-۱- استخراج داده‌های حرکتی از سیگنال گفتار ۶۹
۵-۱-۱ آماده‌سازی و پردازش‌های اولیه ۷۰
۵-۱-۲- روش پیشگویی خطی ۷۰
۵-۱-۳- تخمین سطح مقطع مجرای صوتی ۷۵
۵-۲- شبیه‌سازی مدل پیشنهادی و تحلیل نتایج ۷۸
۵-۲-۱- مدل عصبی عضلانی لوله آکوستیکی ۷۸
۵-۲-۲- بررسی فرایند یادگیری تکلم در مدل پیشنهادی ۸۰
۵-۲-۳- بررسی اغتشاشات خارجی ۸۴
۵-۲-۴- استفاده از مدل برای تحلیل رفتار بیماری‌ها ۸۵
۵-۲-۵- مقایسه مدل پیشنهادی با دیگر مدل‌ها ۸۶
۵-۳- سنتز گفتار با استفاده از حرکات ایجاد شده توسط مدل ۸۷
۵-۳-۱- سنتز آوا ۸۷
۵-۳-۲- سنتز جمله ۸۹
۵-۳- خلاصه مطالب ۹۰
۶- جمع‌بندی و پیشنهادات ۹۳
۶-۱- جمع‌بندی ۹۳
۶-۲- پیشنهادات ۹۵
مراجع ۹۸

فهرست اشکال:

شکل ۱-۱: بلوک دیاگرام مدل پیشنهادی برای کنترل سیستم تولید گفتار بر اساس کنترل پیش‌بین. نواحی که با علامت (*) مشخص شده‌اند مورد نظر این پروژه نبوده و مورد بررسی قرار نمی‌گیرند. ۷
شکل ۲-۱: نواحی مختلف قشر مغز وعملکردهایشان [۱]. ۱۲
شکل ۲-۲: نقشه نواحی عملکردی در قشر مغز به ویژه نواحی ورنیکه و بروکا برای درک کلام و تولید کلمات [۶]. ۱۳
شکل ۲-۳: سازماندهی نواحی ارتباطی پیکری، شنوایی و بینایی و ارتباطاتی که با ناحیه ورنیکه دارند [۶]. ۱۵
شکل ۲-۴: مسیرهای مغزی برای درک کلمه شنیده‌شده و سپس گفتن همان کلمه(در این شکل ناحیه شنوایی مشخص نشده است) [۱]. ۱۸
شکل ۲-۵: مسیرهای مغزی برای درک کلمه نوشته‌شده و سپس بیان آن کلمه [۱]. ۱۹
شکل ۲-۶: برش مقطع ساجیتال از سیستم تولید گفتار در انسان [۴]. ۲۰
شکل ۲-۷: بلوک دیاگرامی از سیستم تولید گفتار در انسان [۴]. ۲۱
شکل ۲-۸: عضلات بین‌دنده‌ای و شکمی‌در حین عمل دم و بازدم [۱]. ۲۲
شکل ۲-۹: نمایی از درون حنجره[۲۸]. ۲۳
شکل ۲-۱۰: تارهای صوتی در وضعیت باز و بسته [۳۲]. ۲۴
شکل ۲-۱۱: ساختار غضروفی حنجره از دو نمای مختلف [۳۲]. ۲۴
شکل ۲-۱۲: ماهیچه‌های مختلف حنجره که با انقباضاتشان باعث تغییر طول تارهای صوتی می‌شوند[۲۷]. ۲۵
شکل۲-۱۳: عضلات مختلف حنجره که به قسمت‌های مختلفی از ساختار غضروفی حنجره متصلند و بر حسب محل تکیه گاه باعث جابجایی تارهای صوتی و یا ساختار غضروفی می‌شوند [۲۷]. ۲۶
شکل ۲-۱۴: عصب واگوس که شاخه‌های مختلفی برای عصب‌دهی به عضلات حنجره و ناحیه حلقی دارد[۱]. ۲۷
شکل ۲-۱۵: عصب زبانی که عضلات زبان را کنترل می‌کند [۱]. ۲۸
شکل ۳-۱: ساختارهای رزونانسی که توسط دانشمند روسی در سال ۱۷۷۹ ارایه گردید[۳۳]. ۳۲
شکل ۳-۲: دستگاهی که Wolfgang برای تولید گفتار طراحی کرد که در موزه مونیخ نگهداری می‌شود [۳۴]. ۳۳
شکل ۳-۳: دستگاه Wheatstone که مدلی کاملتر برای تولید گفتار بود [۴]. ۳۳
شکل ۳-۴: دستگاه الکترونیکی VODOR که در زمان خود کامل‌ترین دستگاه تولید گفتار بود [۴]. ۳۴
شکل ۳-۵: مدلسازی مجرای صوتی به وسیله یک لوله بدون تلفات با سطح متغیر با زمان [۴]. ۳۵
شکل ۳-۶: پاسخ فرکانسی مدل تک لوله‌ای برای مجرای صوتی [۳]. ۳۶
شکل ۳-۷: تقسیم مجرای صوتی به لوله‌های بدون اتلاف با قطر متفاوت [۳]. ۳۷
شکل۳-۸: یک سیکل کامل نوسان تارهای صوتی [۲۵]. ۳۹
شکل ۳-۹: مدل جرم و فنر که برای تارهای صوتی ارایه شده است[۲۵]. ۳۹
شکل ۳-۱۰: مدل دو جرمی برای تارهای صوتی که به وسیله دو جرم و فنر و دمپر مدل می‌شود [۲۵]. ۴۰
شکل ۳-۱۱: مدل سه جرمی‌که برای توصیف دقیقتر رفتار تارهای صوتی ارایه گردیده است [۲۵]. ۴۱
شکل۳-۱۲: نواحی فعال در تکرار کلمه شنیده شده بر اساس مدل ورنیکه و بروکا [۲]. ۴۳
شکل ۳-۱۳: مدل شناختی که ورنیکه و بروکا ارایه کردند و دیگر دانشمندان آن را تکمیل نمودند[۲] ۴۳
شکل۳-۱۴: مدلی شناختی که توسط Patterson & Shewell برای زبان ارایه شد [۲]. ۴۴
شکل۳-۱۵: مدلی که توسطSeidenberg and McClelland در ۱۹۸۹ ارایه شده است [۵]. ۴۵
شکل ۳-۱۶: بلوک دیاگرام مدل DIVA برای دریافت و تولید گفتار [۱۱]. ۴۶

شکل ۴-۱: خروجی‌ها در فاصله زمانی N در آینده، که افق پیش‌بینی نامیده می‌شود، با استفاده از مدل پروسه، در هر مرحله و زمان t، پیش‌بینی می‌شوند [۴۶]. ۵۱
شکل ۴-۲: ساختار اصلی کنترل‌کننده پیش‌بین [۱۴]. ۵۲
شکل ۴-۳: مخچه اطلاعات مختلفی را از قشر مغز می‌گیرد و پردازش کرده و دوباره به قشر مغز باز می‌گرداند [۱]. ۵۶
شکل ۴-۴: بلوک دیاگرام سیستم کنترل تولید گفتار بر اساس آزمایشات و شواهد فیزیولوژیکی. ۵۸
شکل ۴-۵: بلوک دیاگرام مدل پیشنهادی برای کنترل سیستم تولید گفتار بر اساس کنترل پیش‌بین. ۵۸
شکل ۴-۶: لوله آکوستیکی و تقسیم آن به قطعات کوچکتر[۴]. ۶۰
شکل ۴-۷: اتصال دو قطعه و برگشت امواج از محل اتصال[۴]. ۶۰
شکل ۴-۸: تارهای صوتی با نوسان خود باعث تغییر جریان پایای شش‌ها به فلوی پالسی شکل می‌شوند[۴]. ۶۲
شکل ۴-۹: مدل تولید گفتار که بر مبنای لوله آکوستیکی ارایه گردیده است (بازسازی شده از [۳،۴]). ۶۳
شکل ۴-۱۰: تکمیل مدل لوله آکوستیکی بر اساس خواص عضلات و تاثیرات سیگنال‌های عصبی روی آن [۲۵]. ..۶۴
شکل ۴-۱۱: مدل لوله آکوستیکی که به وسیله سیگنال‌های عصبی کنترل می‌شود و تغییرات سطح مقطع آن تابعی از سیگنال‌های عصبی(F_i) است. ۶۵
شکل ۴-۱۲: مدل نهایی برای سیستم تولید گفتار که بر مبنای کنترل پیش‌بین است. در شکل H_i معرف تابع تبدیل عضلات قطعه iام است که باعث تغیی در جابجایی‌های آن می‌شوند و G_i معرف مدل قطعه iام در مخچه($H_i(s)$) می‌باشد. ۶۶
شکل ۵-۱: سیگنال زمانی آوای /a/. این آوا به علت واک دار بودن دارای شکلی تقریبا متناوب است. ۷۲
شکل ۵-۲: دامنه سیگنال /a/ و طیف فرکانسی آن. فرکانس نمونه برداری ۲۲۰۵۰ هرتز است. ۷۲
شکل ۵-۳: دامنه سیگنال /a/ و طیف فرکانسی آن که با استفاده از الگوریتم پیشگویی خطی مرتبه ۱۰ ایجاد شده است. ۷۳
شکل ۵-۴: طیف سیگنال برای درجات مختلف پیشگویی خطی. با افزایش درجه طیف به طیف سیگنال نزدیک‌تر می‌شود. ۷۳
شکل ۵-۵: سیگنال زمانی مربوط به آوای /h/. ۷۴
شکل ۵-۶: مقایسه طیف سیگنال اصلی و سیگنال تقریبی حاصل از روش پیشگویی خطی مرتبه ۱۰. خط قرمز و خط چین سیگنال اصلی و خط پیوسته سیگنال حاصل از تقریب را نشان می‌دهد. ۷۵
شکل ۵-۷: سطح مقطع مجرای صوتی برای قطعات مختلف در طول زمان نشان داده شده است. ۷۶
شکل ۵-۸: جابجایی هر کدام از قطعات در طول زمان. ۷۷
شکل ۵-۹: جابجایی هر کدام از قطعات در طول زمان که به وسیله شماره قاب مشخص شده است. ۷۷
شکل ۵-۱۰: آزمایشی که برای ثبت داده‌های ناشی از تحریک الکتریکی عضلات انجام شده است [۵۷]. ۷۹
شکل ۵-۱۱: پاسخ ضربه(بالا) و پاسخ پله مدل(پایین) حرکتی لوله آکوستیکی. ۸۰
شکل ۵-۱۲: سیستم کنترل وقتی مدل مناسبی نداشته باشد نمی‌تواند به خوبی پروسه را کنترل کند. مدلی که از پروسه در کنترل‌کننده پیش‌بین بکار رفته است یک تابع تبدیل درجه اول است. ۸۱
شکل ۵-۱۳: سیگنال مرجع، داده‌های حرکتی بخش دهم در حین بیان آوای /a/ است. مدل بکار رفته در کنترل‌کننده پیش‌بین یک تابع تبدیل خطی مرتبه است. ۸۲
شکل ۵-۱۴: سیستم کنترل وقتی مدل کامل را داشته باشد می‌تواند به خوبی پروسه را کنترل کند. ۸۳
شکل ۵-۱۵: پس از آنکه مخچه سیستم آرتیکولاتورها به صورت کامل شناسایی کرد و مدل نمود، سیستم کنترلی با استفاده از این مدل به خوبی می‌تواند آرتیکولاتورها را کنترل کند. ۸۳

شکل ۵-۱۶: اغتشاش به صورت یک پالس با دامنه ۱٫۵- سانتی‌متر به صورت جمع شونده به خروجی اعمال می‌شود. همانطور که در شکل مشخص شده است مدل به خوبی تاثیر این اغتشاش را کاهش داده است.....۸۴
شکل ۵-۱۷: مدل کنترلی ارایه شده به خوبی اغتشاشات خارجی(نویز سفید) را حذف می‌کند.....۸۵
شکل ۵-۱۸: سیگنال بازسازی شده و سیگنال حاصل از تغییرات سطح مقطع ناشی از حرکات مدل. سیگنال مورد بررسی آوای /a/ است. ۸۸
شکل ۵-۱۹: سیگنال بازسازی شده و سیگنال حاصل از تغییرات سطح مقطع ناشی از حرکات مدل. سیگنال مورد بررسی آوای /h/ است.....۸۹

فهرست جداول

جدول ۵-۱: مقادیر سطح مقطع محاسبه شده از خروجی مدل و مقادیر اصلی بدست آمده از سیگنال گفتار. ۸۹

فصل اول:

کلیات

۱- مقدمه

یکی از بزرگترین تفاوت‌های بین انسان‌ها و حیوانات قدرت تفکر و توانایی انتقال افکار درونی به همنوعان خود است. انسان‌ها برای انتقال افکار درونی خویش به ایجاد ارتباط با یکدیگر می‌پردازند. محققان این توانایی ذهنی بشر را که به واسطه آن افکار درونی را تبدیل به سیگنال‌هایی کرده که با آن‌ها به ایجاد ارتباط با همنوعان خود می‌پردازد را زبان[۱] می‌نامند [۱]. زبان در واقع شامل ارتباطات صداها و علایم معنادار است که ما را برای توصیف محیط اطراف و افکار درونیمان توانا می‌گرداند [۱،۲]. استفاده از زبان نیازمند هماهنگی بین حافظه، اطلاعات سنسوری و خروجی‌های موتوری است. سه نوع عمده حافظه شنوایی[۲]، بینایی[۳] و معنایی[۴] در زبان مورد استفاده قرار می‌گیرند [۲]. ورودی‌های سنسوری به این حافظه‌ها می‌توانند از طریق پردازش‌های شنوایی، پردازش‌های بینایی و یا پردازش‌های لامسه‌ای منتقل گردند. خروجی‌های موتوری نیز این مفاهیم ذهنی را از طریق گفتار، نوشتار، آواز و یا نقاشی بیان می‌کنند که یا به صورت خودکار، در نتیجه افکار درونی، و یا در اثر پاسخ به خواسته‌های بیرونی به وجود می‌آیند. مهمترین خروجی موتوری زبان، گفتار است. تکلم یک عمل مهارتی است که انسان به مرور زمان یاد می‌گیرد و برای ایجاد ارتباط و انتقال اطلاعات به شنونده بکار می‌برد [۱]. تکلم در انسان با ایده و فکری که گوینده می‌خواهد به شنونده منتقل کند شروع می‌شود. گوینده این فکر و ایده را از طریق یک سری فعالیت‌های نرولوژیکی[۵] و حرکات ماهیچه‌ای به یک موج فشار آکوستیکی تبدیل می‌کند که شنونده با استفاده از سیستم شنوایی خود دریافت کرده و از طریق یک سری فرایندها دوباره به سیگنال نورلوژیکی تبدیل می‌کند [۳].

سیستم تولید گفتار[۶] یکی از پیچیده‌ترین و ناشناخته‌ترین سیستم‌های فیزیولوژیکی است. توصیف رفتار این سیستم و ارایه مدل محاسباتی برای آن می‌تواند کمک شایانی به محققان این زمینه نماید. چرا که با وجود چنین مدلی می‌توان به مطالعه بیماری‌های سیستم تولید گفتار پرداخت و راه‌های درمانی مناسبی برای آن‌ها ارایه نمود.

در این فصل ابتدا اشاره کوتاهی به مدلسازی تولید گفتار خواهد شد و مراکز مغزی مرتبط با گفتار معرفی می‌گردند. در ادامه شواهدی که حاکی از وجود پیش‌بینی در سیستم تولید گفتار است را مرور می‌کنم و سپس به معرفی مدلی که در این پژوهش ارایه می‌گردد، پرداخته می‌شود. در نهایت اهداف پروژه و روش انجام آن و روند تنظیم پایان نامه را ذکر خواهم نمود.

---

[۱] Language
[۲] Auditory
[۳] Visual
[۴] Semantic
[۵] Neurological
[۶] Speech Production

۱-۱- مدلسازی تولید گفتار

از آنجا که گفتار مهمترین راه ارتباطی برای انسان به شمار می‌آید، بسیار مورد مطالعه قرار گرفته شده است. برای تولید گفتار مدل‌های مختلفی ارایه گردیده است. اولین مطالعات علمی برای تولید گفتار مربوط به ۲۰۰ سال پیش می‌باشد [۴]. این مطالعات که توسط یک دانشمند روسی انجام گردید منجر به ساخت وسایلی ابتدایی شد که قادر به تولید چند آوای خاص بودند. پس از ایشان مدل‌های مکانیکی مختلفی برای تولید گفتار ارایه گردید. ولی این مدل‌های مکانیکی همگی نیاز به یک کاربر ماهر داشتند و همچنین دارای عملکرد محدود بودند. با پیشرفت صنعت الکترونیک دانشمندان به استفاده از مدارات الکترونیکی برای ساخت ماشینهای سخنگو روی آوردند [۳،۴]. یکی از اولین ماشین‌های تمام الکترونیکی به وسیله J.Q.Stewart در سال ۱۹۲۰ ساخته شد. پس از آن در سال ۱۹۳۵ ماشین VODOR ساخته شد که یکی از کامل‌ترین ماشین‌های سخنگوی زمان خود بود [۴]. گر چه این ماشین‌ها محدودیت عملکردی ماشین‌های مکانیکی را نداشتند ولی برای کار همچنان به یک کاربر ماهر نیاز داشتند که بتواند به تولید گفتار با آن‌ها بپردازد. با ورود کامپیوترها، پیشرفت زیادی در زمینه تولید گفتار صورت گرفت و کیفیت گفتار تولیدی بسیار بهتر گردید. بر این اساس دانشمندان به سراغ مدل‌های ریاضی رفتند. امروزه مدل‌های ریاضی مختلفی برای تولید گفتار ارایه شده است و پیشرفتهای زیادی در این زمینه صورت گرفته است. با وجود تمام پیشرفتهای صورت گرفته عموم مدل‌های ارایه شده برای تولید گفتار دارای پایه و اساس فیزیولوژیکی نیستند چرا که به تولید گفتار به عنوان یک فعالیت عصبی عضلانی نگاه نمی‌کنند و صرفا آن را از دید تولید یک سیگنال مورد بررسی قرار می‌دهند. بنابراین چنین مدل‌هایی نمی‌توانند رفتار عصبی عضلانی این سیستم را به درستی نشان دهند. علاوه بر این نوع مدل‌ها، نوع دیگری نیز وجود دارد که نگاهی متفاوت به گفتار دارند. همانطور که گفته شد گفتار یکی از راه‌ها و در واقع مهم‌ترین راه خروجی زبان است به همین دلیل عموم مدل‌هایی که برای زبان ارایه شده است از گفتار به عنوان خروجی استفاده کرده‌اند. در این مدل‌ها چگونگی شکل گیری جملات بر اساس مفهوم کلمات و گرامر مدلسازی شده‌اند [۲،۵]. به عبارت دیگر این گونه مدل‌ها بیشتر به مراحل عالی‌تر تولید گفتار از جمله تشکیل و چیدمان کلمات و جملات و درک مفاهیم رشته‌های شنیداری پرداخته‌اند. در این گونه مدل‌ها نیز به تولید گفتار به دید یک عمل عصبی عضلانی پرداخته نمی‌شود و اصولا در این مدل‌ها به تولید حرکت پرداخته نمی‌شود. مدل عصبی عضلانی برای تولید گفتار می‌تواند در درک رفتار حرکتی تولید گفتار کمک فراوانی کند و با استفاده از چنین مدلی بررسی دقیق‌تر بیماری‌های عصبی عضلانی تولید گفتار نیز امکان پذیر می‌شود. بر این اساس در این پروژه سیستم تولید گفتار را با استفاده از دیدگاه عصبی عضلانی مورد بررسی قرار می‌گیرد.

۱-۲- مراکز مغزی مرتبط با تولید گفتار

مراکز مغزی مختلفی در تولید و درک گفتار نقش دارند. قسمت عمده این مراکز در قشر مغز قرار دارند [۱]. هنگامی‌که کلمه‌ای شنیده می‌شود اطلاعات شنیداری پس از دریافت توسط گوش به ناحیه شنوایی[۱] مغز ارسال می‌شوند. این قسمت مسئول درک صداهای شنیده شده است و با حافظه شنوایی که صداهای شنیده شده گذشته را دارد ارتباط نزدیکی دارد. صدای شنیده شده با صداهای حافظه مقایسه شده و بازشناسی می‌شود و سپس به ناحیه ورنیکه[۲] ارسال می‌گردد [۱،۶]. همچنین زمانی که کلمه‌ای را بیان می‌کنیم اطلاعات حسی[۳] مربوط به عضلات مرتبط با تولید گفتار به ناحیه حسی ارسال شده و پس از پردازش به ناحیه ورنیکه ارسال می‌شوند [۱]. ناحیه ورنیکه را مسئول درک و فهم گفتار می‌دانند [۲،۵]. در این ناحیه مفاهیم درک می‌شوند. این ناحیه با ارتباطاتی که با قسمت‌های مختلف مغز دارد اعمال عالی تفکر و درک گفتار را انجام می‌دهد و نهایتا پاسخ مناسب را پردازش کرده و برای تبدیل به یک رشته گفتاری به ناحیه بروکا[۴] ارسال می‌کند [۱،۲،۶]. ناحیه بروکا مسئول چیدمان کلمات است این قسمت با توجه به مفهومی که از ناحیه ورنیکه اعمال می‌شود کلمه مناسب را از حافظه بارگذاری کرده و کلمات را پشت سر هم می‌چیند [۶،۵]. پس از ساخت کلمات و رشته‌های گفتاری توسط بروکا، برای ادا شدن توسط اندام‌های گفتاری به قسمت حرکتی اولیه[۵] از قشر مغز ارسال می‌گردند. ناحیه حرکتی اولیه اطلاعات مختلفی را از بروکا و ناحیه حسی دریافت می‌کند [۲،۵،۶]. همچنین این ناحیه ورودیهایی نیز از مخچه[۶][۵،۶] و ساختارهای زیر قشری نظیر عقده‌های قاعده‌ای[۷] [۷] دریافت می کند و بر اساس تمام این اطلاعات دستورات موتوری[۸] را به عضلات و اندام‌های حرکتی تولید گفتار ارسال می‌کند [۱،۶،۵]. مخچه یکی دیگر از مراکز با اهمیت در گفتار است. این ناحیه در دو سطح مختلف در تولید گفتار شرکت می‌کند [۸]. مخچه با بروکا و ناحیه حرکتی اولیه در ارتباطی دو طرفه است [۹] و همین باعث می‌شود که هم در سطح عالی تولید گفتار و هم در سطح اجرایی آن شرکت کند.

۱-۳- شواهد وجود پیش‌بینی در سیستم تولید گفتار

پیش‌بینی در سیستم حرکتی مربوط به عضلاتی چون دست و پا، امری کاملا شناخته شده است و محققان زیادی این موضوع را بررسی کرده‌اند. مدل‌های حرکتی مختلفی نیز برای سیستم

---

[۱] Auditory Cortex
[۲] Wernicke's Area
[۳] Somatosensory
[۴] Broca’s area
[۵] Primary motor cortex
[۶] Cerebellum
[۷] Basal Ganglia
[۸] Motor Command

حرکتی ارایه شده است که هر کدام به نحوی پیش‌بینی را در آن در نظر گرفته‌اند [۱۰]. لیکن همان‌گونه که گفته شد مدل‌هایی که به جنبه حرکتی تولید گفتار پرداخته باشند بسیار کم هستند.

آزمایشاتی که بر روی سیستم تولید گفتار انجام شده است نشان می‌دهند که سیستم کنترل حرکات تولید گفتار نیز مشابه سایر حرکات دارای قابلیت پیش‌بینی اتفاقات آینده است [۹،۱۱،۱۲]. محققان با وارد کردن اغتشاشات متناوب به سیستم تولید گفتار نشان داده‌اند که شخص پس از مدتی قادر به پیش‌بینی اغتشاشات قبل از زمان وقوعشان است و با حرکت اندامهای حرکتی مسیر تولید گفتار این اغتشاشات را حذف می‌کند [۱۱]. علاوه بر این، گزارش‌های گفتار درمانی حاکی از آن است که اشخاصی که برای مدتی از یک پروتز دندانی استفاده می‌کنند در ابتدا نمی‌توانند به مانند قبل تکلم کنند ولی پس از مدتی یاد می‌گیرند بر این مشکل فایق آیند. همچنین اشخاصی که پروتز دندان را خارج می‌کنند در اوایل دچار مشکل در تکلم می‌شوند که بعد از مدتی بهبود می‌یابند [۱۲]. این گونه مشاهدات و همچنین آزمایشات دیگری از این دست اشاره به وجود مدل داخلی در سیستم تولید گفتار دارد. به عبارت دیگر سیستم کنترل حرکات گفتاری دارای مدلی داخلی[۱] از اندام‌های خود است که با استفاده از این مدل اطلاعات مختلفی را در مورد نتایج دستوراتی که به اندام‌ها ارسال می‌شود را قبل از ارسال تخمین می‌زند و از این مدل برای پیش‌بینی و کنترل و جبرانسازی سیستم تولید گفتار استفاده می‌کند [۱۳]. ایده استفاده از کنترل مبتنی بر مدل در صنعت به خوبی بکار گرفته شده است و امروزه کنترلرهای صنعتی زیادی بر این مبنا کار می‌کنند [۱۴]. بر اساس مزایای کنترل مبتنی بر مدل محققان روش‌های کنترلی مختلف مبتنی بر مدل را به سیستم کنترل حرکات بدن نسبت داده‌اند. محققان با طرح آزمایشات مختلف به بررسی وجود مدل‌های داخلی برای گفتار پرداخته‌اند و به نتایج مثبتی در این زمینه رسیده‌اند ولی پیچیدگی ذاتی این سیستم باعث شده است که پیشرفت‌های زیادی در این حوزه صورت نگیرد [۱۳].

بنابراین با توجه به این دو مجموعه آزمایشات و شواهد که نشان دهنده وجود یک مدل داخلی از اعضای حرکتی تولید گفتار در مغز و نیز قابلیت پیش‌بینی نتایج حرکت آرتیکولاتورها[۲]، قبل از انجام حرکت می‌باشند، می‌توان چنین پنداشت که سیستم کنترل تولید گفتار احتمالا مشابه یک کنترل‌کننده پیش‌بین مبتنی بر مدل[۳] عمل می‌کند. شایان ذکر است که محققان نشان داده‌اند که عملکرد سیستم کنترلی بدن در تنظیم امپدانس[۴] عضلات و همچنین تولید مسیر مشابه عملکرد این نوع کنترل‌کننده‌ها در کنترل پروسه‌ها است [۱۰،۱۵] لذا منطقی به نظر می‌رسد که مشابه این نوع کنترل‌کننده‌ها برای سیستم تولید گفتار هم بکار رود.

---

[۱] Internal Model
[۲] Articulator
[۳] Model Predictive Controller
[۴] Impedance

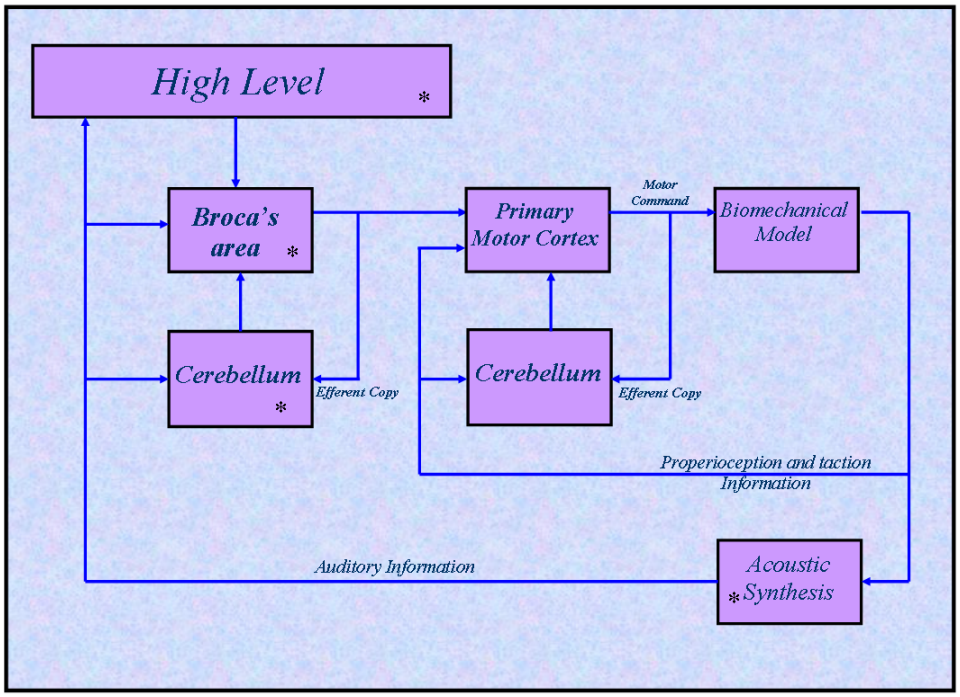
۱-۴- مدل پیشنهادی برای این پروژه

براساس مزایای کنترل‌کننده پیش‌بین مبتنی بر مدل و نیز شواهد وجود مدل داخلی برای سیستم تولید گفتار و نیز نتایج آزمایشاتی که حاکی از وجود پیش‌بینی در سیستم تولید گفتار است، در این پروژه از ایده کنترل پیش‌بین استفاده می‌شود و مدلی برای سیستم عصبی عضلانی تولید گفتار بر همین مبنا ارایه می‌گردد. هر کنترل‌کننده پیش‌بین شامل چند قسمت اصلی است که می‌توان آن‌ها را به شکل زیر خلاصه کرد [۱۴]:

- فیدبک(اطلاعات گذشته سیستم)
- مدل پروسه(که شامل مدل پروسه و مدل اغتشاشات)
- بهینه‌سازی[۱] (که برای آن روش‌های مختلفی وجود دارد)

در سیستم تولید گفتار همانطور که قبلا هم گفته شد دو دسته اطلاعات سنسوری داریم. دسته اول سنسورهای حسی و لامسه است که اطلاعات مربوط به وضعیت و موقعیت عضلات و آرتیکولاتورها را شامل می‌شوند. دسته دوم سنسورهای شنوایی است که اطلاعات شنیداری را به مغز منتقل می‌کنند. مدل یکی دیگر از اجزای اصلی کنترل پیش‌بین است. شواهد و مطالعاتی که بر روی مخچه انجام شده است نشان می‌دهد که مخچه نقش یک مدل را دارد [۱۷]. این گونه مطالعات بر روی مباحث حرکت اندامهایی چون دست یا پا صورت گرفته‌اند و مطالعات کمی‌در مورد اندامهای گفتاری صورت گرفته است. همانطور که گفته شد در سیستم تولید گفتار دو دسته اطلاعات سنسوری وجود دارد که هرکدام نقش خاصی را بر عهده دارند. یک کپی از این اطلاعات هر دو دسته قبل از ورود به مغز به مخچه نیز ارسال می‌گردد [۸،۹،۱۸]. مخچه علاوه بر این ورودی‌هایی که بیان شد خروجی‌های مختلفی نیز به ناحیه حرکتی اولیه و بروکا و نیز نواحی دیگری مانند ورنیکه دارد [۹،۱۸،۲۱]. لذا می‌توان چنین گفت که مخچه علاوه بر مدلسازی سیستم تولید گفتار نتایج پیش‌بینی بر اساس مدل را به ناحیه حرکتی اولیه ارسال می‌کند. مشابه این حالت برای ارتباطات بروکا و مخچه نیز وجود دارد. بهینه‌سازی قسمت دیگری از یک کنترل‌کننده پیش‌بین است. نواحی حرکتی اولیه و بروکا با گرفتن اطلاعات مرجع از مراکز بالاتر و نیز اطلاعات سنسوری برگشتی از طرف اندام‌ها و همچنین اطلاعات مربوط به پیش‌بینی توسط مدل داخلی که در مخچه قرار دارد دستورات ارسالی به اندام‌ها را بهینه می‌کنند. بر اساس مطالب بیان شده تناظرات بین سیستم کنترلی تولید گفتار و کنترل پیش‌بین کامل می‌شوند. در شکل ۱-۱ بلوک دیاگرام مربوط به مدل پیشنهادی برای سیستم کنترلی تولید گفتار که بر اساس ارتباطات نواحی مختلف مغز که در کنترل سیستم تولید گفتار نقش دارند و نیز تناظرات بین این ارتباطات و کنترل پیش‌بین ارایه شده است.

[۱] Optimization



شکل ۱-۱: بلوک دیاگرام مدل پیشنهادی برای کنترل سیستم تولید گفتار بر اساس کنترل پیش‌بین. نواحی که با علامت (*) مشخص شده‌اند مورد نظر این پروژه نبوده و مورد بررسی قرار نمی‌گیرند.

۱-۵- اهداف پروژه و روش انجام آن

۱-۵-۱- اهداف پروژه

هدف اصلی در این پژوهش، مطالعه سیستم تولید گفتار در سطح حرکتی و مدلسازی جنبه عصبی عضلانی تولید گفتار است. از دیگر اهداف این پروژه، شناخت مراکز مغزی مرتبط با تولید گفتار و چگونگی ارتباطات بین آن‌ها است. در نهایت در این پروژه یک مدل بر اساس تناظرات بین فعالیت‌های مراکز مغزی و کنترل‌کننده پیش‌بین مبتنی بر مدل ارایه می‌گردد که رفتار عصبی عضلانی سیستم تولید گفتار را در سطح حرکتی و با استفاده از فیدبک حسی مدلسازی می‌نماید. به عبارت دیگر هدف از انجام این پروژه ارایه مدلی محاسباتی برای سیستم عصبی عضلانی تولید گفتار است که با استفاده از آن بتوان رفتار اسیستم تولید گفتار را مورد بررسی قرار داده و نقش هر یک از نواحی مغزی مرتبط با تولید گفتار به خوبی شناخته شود. چنین مدل‌هایی که بر پایه واقعیت عصبی عضلانی سیستم تولید گفتار می‌باشند می‌توانند در بهبود روش‌های تشخیص و درمان بیماری‌های عصبی عضلانی و یا حتی بیماری‌های مرتبط با سطوح بالای تولید گفتار نظیر لکنت نیز مفید واقع شوند.

۱-۵-۲- فرضیات مدلسازی

در زیر به فرض‌های اعمال شده برای ارایه مدل پیشنهادی و نیز پیاده‌سازی آن، اشاره می‌شود.

۱. در این پروژه بر روی تولید گفتار در اشخاصی که تکلم را کاملا یاد گرفته‌اند بحث می‌شود. به عبارت دیگر فرض می‌شود مدل‌هایی که در مخچه برای اندام‌ها بکار می‌رود قبلا با تمرین و تجربه بدست آمده‌اند و الگوریتم مربوط به فرایند یادگیری مورد بررسی قرار نمی‌گیرد.
۲. همانطور که گفته شد تولید گفتار در سطوح مختلفی انجام می‌شود. در این پروژه تولید حرکات مجرای صوتی با استفاده از فیدبک حسی مورد بررسی قرار می‌گیرد و به سطوح بالاتری نظیر شنوایی پرداخته نمی‌شود.
۳. برای ممانعت از پیچیدگی مدل از نقش عقده‌های قاعده‌ای صرفنظر می‌شود.
۴. در این پروژه تولید گفتار که تحت تاثیر وضعیت روحی روانی و احساسات باشد مد نظر نیست.

۱-۵-۳- استخراج داده‌ها

در این پروژه از مجموعه دادگان فارس دات استفاده شده است. برای تخمین جابجایی از روی سیگنال گفتار ابتدا باید پردازش‌های اولیه‌ای را بر روی آن انجام داد. به دلیل غیر ایستا بودن سیگنال گفتار و نغییر خواص آن در طول زمان، برای سیگنال گفتار از تحلیل‌های زمان کوتاه استفاده می‌شود با استفاده از الگوریتم پیشگویی خطی[۱] به تخمین سطح مقطع مجرای صوتی پرداخته می‌شود.

۱-۵-۴- پیاده سازی کامپیوتری

برای پیاده سازی الگوریتم مربوط به استخراج داده‌های حرکتی و نیز پیاده‌سازی مدل پیشنهادی از نرم افزار مطلب[۲] استفاده شده است. برای پیاده‌سازی از دو جعبه ابزار سیمولینک[۳] و کنترل پیش‌بین از این نرم افزار استفاده شده است. استخراج داده‌ها نیز با کمک جعبه ابزار پردازش سیگنال[۴] میسر شده است.

---
[۱] Linear Predictive
[۲] MATLAB
[۳] Simulink
[۴] Signal Processing

۱-۶- روند تنظیم پایان نامه

موضوع اصلی این تحقیق، ارائه مدلی عصبی عضلانی برای تولید گفتار است. برای این منظور در فصل دوم فیزیولوژی سیستم تولید گفتار را که شامل دو قسمت عمده دستگاه عصبی و اندام‌های اجرایی تولید گفتار هستند، معرفی می‌شوند. در فصل سوم به بیان دیدگاه‌هایی که سیستم تولید گفتار از منظر آن‌ها مورد مدلسازی قرار گرفته است پرداخته می‌شود. در فصل چهارم مدل پیشنهادی در این پروژه معرفی می‌شود. در این فصل ابتدا مقدمه‌ای بر کنترل پیش‌بین آورد شده است. سپس شواهد و آزمایشاتی که نشان دهنده وجود پیش‌بینی و مدل داخلی در سیستم تولید گفتار هستند، ارایه می‌گردد. در ادامه مدل کنترلی ارایه شده معرفی می‌شود و بعد از آن مدل آرتیکولاتورها که برای شبیه‌سازی مورد استفاده قرار می‌گیرد معرفی می‌شود. در فصل پنجم نتایج شبیه‌سازی آورده شده است که این فصل شامل سه دسته شبیه‌سازی است. دسته اول نتایج شبیه‌سازی مربوط به استخراج داده‌های حرکتی از سیگنال گفتار را نشان می‌دهد. دسته دوم، نتایج شبیه‌سازی مدل است که بر اساس داده‌های حرکتی به کنترل مدل آرتیکولاتورها می‌پردازد. در نهایت، دسته سوم بازسازی سیگنال گفتار از روی حرکات کنترل شده توسط مدل پیشنهادی است. در فصل ششم از مطالب گفته شده جمع‌بندی به عمل می‌آید و در انتها پیشنهاداتی برای ادامه تحقیقات در این زمینه آورده شده است.

فصل دوم:

فیزیولوژی سیستم تولید گفتار

۲- فیزیولوژی سیستم تولید گفتار

گفتار به عنوان مهم‌ترین ابزار ارتباطی برای انسان‌ها تلقی می‌گردد. سیستم تولید گفتار یکی از پیچیده‌ترین و ناشناخته‌ترین سیستم‌های فیزیولوژیکی است. یک گوینده با حرکت دادن و تغییر مکان اندام‌هایی که در مسیر هوایی از حنجره تا لب‌ها هستند، باعث تغییر فشار و جریان هوای داخل مسیر هوایی می‌شود. این تغییرات فشار و جریان هوا منجر به تولید سیگنال‌های آکوستیکی می‌گردند که قابل شنیدن هستند. برای تولید این سیگنال‌های قابل شنیدن که در واقع اطلاعاتی از جانب گوینده هستند، باید قسمت‌های مختلفی در مسیر هوایی به صورت هماهنگ حرکت کنند.

مشابه سایر سیستم‌های حرکتی در انسان، تولید گفتار نیز نتیجه برهمکنش و هماهنگی بین اعصاب به عنوان کنترل‌کننده و عضلات به عنوان عملگر می‌باشد. از آنجا که مدلسازی و درک رفتار این سیستم نیازمند شناخت کامل آن می‌باشد در این فصل ابتدا مراکز مغزی و اعصاب مرتبط با تولید گفتار را مورد بررسی قرار می‌گیرند و سپس به بررسی اندامهای مسیر صوتی از ششها تا لب‌ها، پرداخته می‌شود.

۲-۱- مراکز مغزی مرتبط با گفتار

نواحی مختلفی از مغز بر روی تولید و درک گفتار نظارت می‌کنند و هر کدام جنبه خاصی از این فرایند بسیار پیچیده را تحت کنترل دارند. نواحی مختلفی در قشر مغز و همچنین ساختارهای زیر قشری به همراه مخچه در تولید و درک گفتار شرکت دارند.

۲-۱-۱- نواحی قشری مرتبط با گفتار

مطالعاتی که جراحان، نورولوژیستها و فیزیولوژیست‌ها بر روی انسان انجام داده‌اند نشان می‌دهد که نواحی مختلف قشر مغز دارای عملکردهای جداگانه‌ای هستند. محققان با استفاده از تحریک الکتریکی قشر مغز در بیماران حین معاینات عصبی قبل و بعد از برداشتن قسمت‌هایی از قشر مغز بسیاری از این عملکردها را مشخص کرده‌اند [۱،۶]. بیمارانی که به طور الکتریکی تحریک می‌شوند افکار ناشی از تحریک را بازگو می‌کنند و گاهی حرکاتی را نیز انجام می‌دهند و صدا یا حتی کلمه‌ای را به طور خودبخودی بیان می‌کنند. دانشمندان با کنار هم قرار دادن اطلاعات بدست آمده از منابع مختلف نقشه جامعی از عملکرد قسمت‌های مختلف قشر مغز بدست آورده‌اند [۶]. در شکل ۲-۱ بر همین مبنا نواحی اولیه و ثانویه قشر حرکتی[۱] و همچنین نواحی اصلی اولیه و ثانویه

[۱] Motor Cortex