بسم الله الرحمن الرحيم

١٥٧٤٠

دانشکده فنی مهندسی

گروه مهندسی برق

# طراحی و پیاده سازی یک تراشه شبکه عصبی برای کاربرد شناسایی گفتار با استفاده از تکنولوژی CMOS

مهدی توکلی

پایان نامه جهت دریافت درجه کارشناسی ارشد

اساتید راهنما :

دکتر بهبود مشعوفی

پرفسور عبدالله خویی

۱۳۸۸

وزارت علوم، تحقیقات و فناوری
پژوهشگاه علوم و فناوری اطلاعات ایران
IRANDOC
مرکز اطلاعات و مدارک علمی ایران

۱۵۷۴۶۰

۵ /۳/ ۱۳۹۰

**تشکر و قدردانی:**

اکنون که به حول و قوه ی الهی این پایان نامه به پایان رسیده است بر خود لازم می دانم از زحمات آقای دکتر بهبود مشعوفی که در طی انجام این پایان نامه با نظرات و راهنمایی های خود مرا یاری نمودند کمال تشکر را داشته باشم.همچنین لازم است از آقای دکتر خویی که با نظرات مفید خود باعث پربارتر شدن این پایان نامه شدند قدردانی نمایم . از کلیه عزیزانی که در رابطه با این پایان نامه متقبل زحمت گشته اند خصوصا از پدر و مادرم که در طول سال های تحصیل همواره پشتیبان من بوده اند وهمچنین همسر عزیزم که همپای من زحمات و سختی های انجام این پایان نامه را به جان خرید و همچنین برادر عزیزم آقای مهندس هادی توکلی کمال تشکر و قدردانی را دارم.

دانشگاه ارومیه

پایان نامهٔ کارشناسی ارشد الکترونیک خانم / آقای مهدی ترکمن به تاریخ ۸۸/۱۱/۱۵ شماره

مورد پذیرش هیأت محترم داوران با رتبه بسیار خوب و نمرهٔ ۱۶ قرار گرفت.

۱- استاد راهنما و رئیس هیأت داوران: دکتر مهدی مشوق ، دکتر حزبی

۲- استاد مشاور: دکتر حزبی

۳- داور خارجی دکتر جهل امرانی

۴- داور داخلی دکتر حدیدی

۵- نماینده تحصیلات تکمیلی : دکتر نوری نیا

حق طبع و نشر مطالب این پایان نامه
در انحصار دانشگاه ارومیه می باشد

**فهرست**

**فصل اول : آشنایی مقدماتی با علم آواشناسی ، صوت و واج های زبان فارسی**


۱-۱ مقدمه ..... ۱

۱-۲ مکانیزم تولید صحبت ..... ۱

۱-۳ اندام های واک ساز ..... ۱

۱-٤ اندام های گویایی ..... ۳

۱-۵ آواشناسی زبان فارس..... ٤

۱-۵-۱ آوایی وغیر آوایی ..... ٤

۱-۵-۲ همخوان ها و واکه ها ..... ٤

۱-۵-۳ واج و واج گونه ..... ٦

۱-٦ مدل نظری سیستم گفتار انسان ..... ۷

۱-۷ خلاصه ..... ۸


**فصل دوم : پردازش دیجیتال سیگنال صوت**


۲-۱ مقدمه ..... ۹

۲-۲ آنالیز سیگنال صحبت در حوزه ی زمان ..... ۱۰

۲-۲-۱ آنالیز سیگنال صحبت در حوزه ی زمان با استفاده از انرژی در دوره ی زمانی کوتا..... ۱۱

۲-۲-۲ آنالیز سیگنال صحبت در حوزه ی زمان با استفاده از متوسط نرخ عبور از صفر در دوره ی زمانی کوتاه ..... ۱۲

۲-۲-۳ آنالیز سیگنال صحبت در حوزه ی زمان با استفاده از تابع خود همبستگی در دوره ی زمانی کوتاه. .....۱۳

۲-۳ آنالیز سیگنال صحبت در حوزه ی فرکانس ..... ۱٤

۲-۳-۱ آنالیز سیگنال صحبت در حوزه ی فرکانس با استفاده از پوش طیف ..... ۱٤

۲-۳-۲ آنالیز سیگنال صحبت در حوزه ی فرکانس با استفاده از روش تخمین طیف توان ..... ۱۵

۲-۳-۳ آنالیز سیگنال صحبت با استفاده از ضرایت کپسترال ..... ۱۷

٢-٤ تحلیل سیگنال صوت با استفاده از پیشگویی خطی ..... ١٨
١-٤-٢ مدل بندی سیستم به روش تمام قطب ..... ١٩
٢-٤-٢ پیشگویی خطی مستقیم ..... ٢٠
٣-٤-٢ مرتبه ی پیشگویی خطی ..... ٢٣
٢-٥ خلاصه ..... ٢٤

**فصل سوم : مقدمه ای بر شبکه های عصبی و آشنایی با شبکه ی عصبی TDNN**

٣-١ مقدمه ..... ٢٦
٣-٢ آشنایی با شبکه های عصبی ..... ٢٧
٣-٣ ویژگی های شبکه های عصبی ..... ٢٩
١-٣-٣ حافظه ی انجمنی ..... ٣٠
٢-٣-٣ دسته بندی ..... ٣٠
٣-٣-٣ قدرت تعمیم ..... ٣١
٣-٤ شبکه ی MLP ..... ٣٢
٣-٥ الگوریتم آموزشی BP ..... ٣٣
٣-٦ شبکه ی عصبی TDNN ..... ٣٦
٣-٧ خلاصه ..... ٣٩

**فصل چهارم : شناسایی فونم های زبان فارسی با استفاده از شبکه ی عصبی TDNN**

٤-١ مروری بر کارهای انجام شده ..... ٤٠
٤-٢ تعیین بردار ورودی شبکه ی عصبی ..... ٤٥
٤-٣ پیش پردازش سیگنال صوت ..... ٤٦
١-٣-٤ فیلتر پایین گذر و مبدل آنالوگ به دیجیتال ..... ٤٧

٤-٣-٢ فیلتر دیجیتال پیش تاکید ..... ٤٧
٤-٣-٣ تشخیص ابتدا و انتهای فونم ..... ٤٧
٤-٣-٤ قاب بندی ..... ٤٩
٣-٣-٥ پنجره گذاری..... ٥٠
٤-٤ استخراج بردار ویژگی ها..... ٥٠
٤-٥ محاسبه ی ضرایت کپسترال از روی طیف حاصل از LPC ..... ٥٤
٤-٥-١ وزن دادن به ضرایت کپسترال ..... ٥٥
٤-٥-٢ محاسبه ی ضرایت دلتا کپسترال ..... ٥٦
٤-٦ تعیین مرتبه ی تحلیل پیشگویی خطی ..... ٥٨
٤-٧ ارائه نتایج شبیه سازی انجام شده توسط نرم افزار MATLAb ..... ٦١
٤-٨ خلاصه ..... ٦٩

**فصل پنجم :پیاده سازی سخت افزاری شبکه ی عصبی TDNN**

٥-١ مقدمه ..... ٧٠
٥-٢ تعیین تعداد بیت های اعداد در محاسبات شبکه و همچنین نوع اعداد از لحاظ fixed point
یا floating point ..... ٧١
٥-٣ بررسی بلوک های استفاده شده برای پیاده سازی شبکه ی عصبی TDNN ..... ٧٢
٥-٣-١ پیاده سازی تابع فعالیت نرون(تابع سیگموئید) ..... ٧٣
٥-٣-٢ پیاده سازی بلوک ضرب کننده ..... ٨٠
٥-٣-٣ پیاده سازی بلوک جمع کننده ..... ٨٨
٥-٤ پیاده سازی شبکه ی عصبی TDNN ..... ٩٥
٥-٤-١ پیاده سازی شبکه ی عصبی TDNN توسط تکنولوژی CMOS..... ٩٩
٥-٤-٢ پیاده سازی شبکه ی عصبی TDNN به زبان VHDL (FPGA) ..... ١٠١

۵-۵ مشخصات چیپ طراحی شده ..... ۱۰۴

۵-۶ مقایسه با کار دیگران ..... ۱۰۵

۵-۷ نتیجه گیری و پیشنهادات ..... ۱۰۷

**منابع و مراجع** ..... ۱۰۹

## فهرست اشکال

### فصل اول :


شکل ۱-۱ اندام گویایی انسان ..... ۲

شکل ۲-۱ سیگنال زمانی واج /e/ ..... ۳

شکل ۳-۱ سیگنال زمانی واج /س/ ..... ۳

شکل ٤-۱ مدل سیستم صوتی انسان ..... ۸


### فصل دوم :


شکل ۱-۲ شکل موج زمانی یک سیگنال نمونه ..... ۱۱

شکل ۲-۲ شکل موج یک سیگنال به همراه نمودار انرژی آن ..... ۱۲

شکل ۳-۲ استفاده از دو روش اندازه گیری انرژی و نرخ عبور از صفر برای تعیین محدوده کلمه ...۱۳

شکل ٤-۲ الف) شکل موج زمانی حرف صدا دار /آ/ ب) طیف فرکانسی حرف صدادار /آ/ .....۱٦

شکل ٥-۲ مدل دستگاه صوتی انسان ..... ۲۰

شکل ٦-۲ اثر افزایش مرتبه ی پیشگویی در تطبیق فرکانسی ..... ۲۳

شکل ۷-۲ اثر افزایش مرتبه ی پیشگویی در خطای پیشگویی ..... ۲٤


### فصل سوم :


شکل ۱-۳ مدل ساده ای از یک نرون عصبی مصنوعی ..... ۲۸

شکل ۲-۳ انواع توابع تصمیم گیری مورد استفاده در نرون ..... ۲۸

شکل ۳-۳ ساختار عمومی یک شبکه عصبی MLP ..... ۳۳

شکل ٤-۳ شبکه ی عصبی چند لایه ی پیش خور ..... ۳٤

شکل ٥-۳ ساختار عمومی یک شبکه ی TDNN با استفاده از یک شبکه ی MLP ..... ۳۸

شکل ٦-۳ شبکه ی عصبی TDNN ..... ۳۹


### فصل چهارم :


شکل ۱-٤ بلوک دیاگرام MAC مدار کلسمن ..... ٤۳

شکل ۲-٤ بلوک دیاگرام نروچی کلسمن ..... ٤٤

شکل ۳-٤ بلوک دیاگرام استخراج ویژگی ها ..... ٤٦
شکل ٤-٤ بلوک دیاگرام واحد پیش پردازش سیگنال صوت ..... ٤٧
شکل ٥-٤ قسمتی از یک شکل موج قبل از اعمال فیلتر ..... ٤٨
شکل ٦-٤ سیگنال شکل ٤-٣ بعد از اعمال فیلتر ..... ٤٨
شکل ٧-٤ سیگنال زمانی کلمه ی / پنج/ ..... ٤٩
شکل ٤-٨ یک فریم ٢٥٦ نمونه ای قبل از اعمال پنجره ..... ٥١
شکل ٩-٤ نمونه های شکل ٤-٦ بعد از اعمال پنجره ی همینگ ..... ٥١
شکل ١٠-٤ ساختمان یک فیلتر IIR ..... ٥٣
شکل ١١-٤ تقریب مجرای صوتی انسان با استفاده از لوله های صوتی ..... ٥٣
شکل ١٢-٤ تاثیر مرتبه ی پیشگویی در تغییرات خطای پیشگویی ..... ٥٨
شکل ١٣-٤ تاثیر تعداد نرون های لایه ی میانی در دقت شبکه ..... ٦٠
شکل ١٤-٤ نمودار آموزش شبکه ی MLP برای شناسایی فونم ها ی زبان فارسی ..... ٦٢
شکل ١٥-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم a ..... ٦٢
شکل ١٦-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم e ..... ٦٣
شکل ١٧-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم u ..... ٦٣
شکل ١٨-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم o ..... ٦٤
شکل ١٩-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم i ..... ٦٤
شکل ٢٠-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم y ..... ٦٥
شکل ٢١-٤ نمودار آموزش شبکه ی TDNN برای شناسایی فونم های زبان فارسی ..... ٦٥
شکل ٢٢-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم a ..... ٦٦
شکل ٢٣-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم e ..... ٦٦
شکل ٢٤-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم u ..... ٦٧
شکل ٢٥-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم o ..... ٦٧
شکل ٢٦-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم l ..... ٦٨
شکل ٢٧-٤ پاسخ شبکه ی MLP به ازای اعمال الگوی فونم y ..... ٦٨

**فصل پنجم:**

شکل ١-٥ نمودار تابع سیگموئید ..... ٧٤

شکل ۲-۵ بلوک دیاگرام تابع سیگموئید ..... ۷۶
شکل ۳-۵ پیاده سازی تابع سیگموئید با استفاده از LUT ..... ۷۷
شکل ٤-۵ شبیه سازی تابع سیگموئید با استفاده از نرم افزار Hspise ..... ۸۰
شکل ۵-۵ بلوک دیاگرام ضرب کننده بوث ۸*۸ ..... ۸۲
شکل ٦-۵ مدار کد کننده بوث ..... ۸۳
شکل ۷-۵ زیر مدار انتخاب کننده ی بوث..... ۸٤
شکل ۸-۵ چگونگی تولید حاصل ضرب های جزئی با استفاده از الگوریتم بوث ..... ۸۵
شکل ۹-۵ شبیه سازی بلوک ضرب کننده ..... ۸۸
شکل ۱۰-۵ بلوک دیاگرام جمع کننده ٤ بیتی با پیش بینی کننده بیت نقلی ..... ۸۹
شکل ۱۱-۵ شماتیک مدار Full adder ..... ۸۹
شکل ۱۲-۵ DFF استفاده شده برای پیاده سازی رجیستر ۱٦ بیتی ..... ۹۰
شکل ۱۳-۵ شبیه سازی بلوک MAC ..... ۹۳
شکل ۱٤-۵ شبیه سازی یک نرون نمونه..... ۹۵
شکل ۱۵-۵ بلوک دیاگرام سخت افزاری شبکه ی پیاده سازی شده ..... ۹۷
شکل ۱٦-۵ مدار داخلی شمارنده ..... ۹۸
شکل ۱۷-۵ پیاده سازی شبکه ی عصبی TDNN با استفاده از تکنولوژی CMOS ..... ۱۰۱
شکل ۱۸-۵ پیاده سازی شبکه ی عصبی TDNN توسط FPGA ..... ۱۰۳
شکل ۱۹-۵ Layout شبکه ی عصبی TDNN ..... ۱۰٤

**فهرست جداول :**

**فصل اول :**

جدول ۱-۱ واجگونه های همخوانی زبان فارسی ..... ۵

جدول ۱-۲ واجگونه های واکه ای زبان فارسی ..... ٦

**فصل دوم :**

جدول ۱-۲ محدوده فرکانس فرمنت برای افراد مونث و مذکر ..... ۱۷

**فصل چهارم :**

جدول ۱-٤ دقت شناسایی شبکه عصبی TDNN برای شناسایی فونم های زبان فارسی بر حسب تاخیر موجود در لایه ورودی ..... ٦۱

**فصل پنجم:**

جدول ۱-۵ اعداد ذخیره شده در ROM ..... ۷۷

جدول ۲-۵ ورودی- خروجی تابع سیگموئید ..... ۷۸

جدول ۳-۵ خروجی بلوک اینکدر برحسب ورودی آن ..... ۸۱

جدول ٤-۵ ورودی – خروجی بلوک ضرب کننده ..... ۸۵

جدول ۵-۵ ورودی – خرروجی بلوک MAC ..... ۹۱

جدول ٦-۵ ورودی- خروجی نرون شبیه سازی شده ..... ۹۳

جدول ۷-۵ دقت شبکه طراحی شده در شناسایی فنوم های زبان فارسی و مقایسه با کار دیگران ..... ۱۰۵

جدول ۸-۵ مقایسه ی نرو چیپ طراحی شده با کار های دیگران ..... ۱۰٦

جدول ۹-۵ نتایج پیاده سازی شبکه ی TDNN توسط FPGA و مقایسه با کار دیگران ..... ۱۰۷

# فصل اول :

# آشنایی مقدماتی با علم آواشناسی،صوت وواج های زبان فارسی

## ۱-۱مقدمه:

برای پردازش سیگنال گفتار ابتدا باید با ماهیت و ویژگی سیگنال صوت آشنا شویم . در این فصل به صورت اجمالی و در حد نیاز چگونگی تولید سیگنال صوت و ویژگی‌های گفتار را مورد بررسی قرار می دهیم و با اصطلاحاتی که در حوزه پردازش گفتار مطرح است آشنا می شویم .

در بخش اول این فصل چگونگی تولید سیگنال صوت و اعضا تولید کننده سیگنال صوت را مورد بررسی قرار داده و مدل ارائه شده برای دستگاه صوتی را توضیح می دهیم. در بخش دوم ماهیت و ویژگی سیگنال صوت را مورد بررسی قرار داده و با اصطلاحات مطرح در حوزه آواشناسی آشنا می شویم و در پایان سیگنال‌های گفتار زبان فارسی را مورد بررسی قرار خواهیم داد .

## ۱-۲ مکانیزم تولید صحبت :

اندام‌های گفتاری وظیفه تولید صحبت را بر عهده دارند . این اندام ها که اندام های تنفسی یا گوارشی هستند ، به عنوان وظیفه دوم خود ، نقش تولید صحبت را بر عهده دارند . این اندام ها که در تولید گفتار به کار می روند ، بر اساس نوع تحریک به سه دسته واکساز ، تنفسی و گویایی تقسیم می شوند . شکل (۱-۱) اندام های گویایی را نشان می‌دهد .

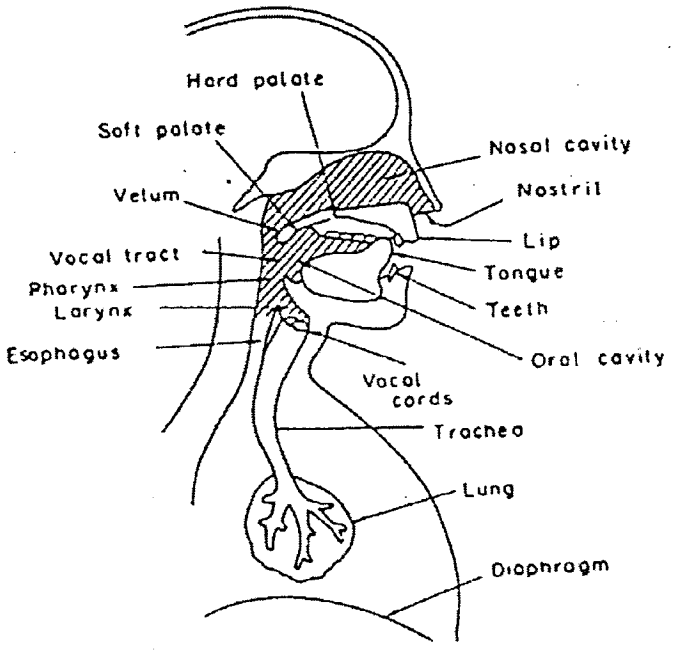
## ۱-۳ اندام‌های واک‌ساز :

این اندام ها در محفظه ی غضروفی حلق[۱] قرار گرفته اند . در آخرین حلقه ی فوقانی نای یک حلقه ی کامل وجود دارد که از دیگر حلقه ها سخت تر و سنگین تر می باشد و طول آن نیز بیشتر است. این حلقه کری کوئید[۲] نام دارد و پایه یا قاعده حنجره است. تیروئید[۳] مانند سپر نقش حفاظت و نگهداری از کری‌کوئید و تارهای صوتی را بر عهده دارد و در روی آنها

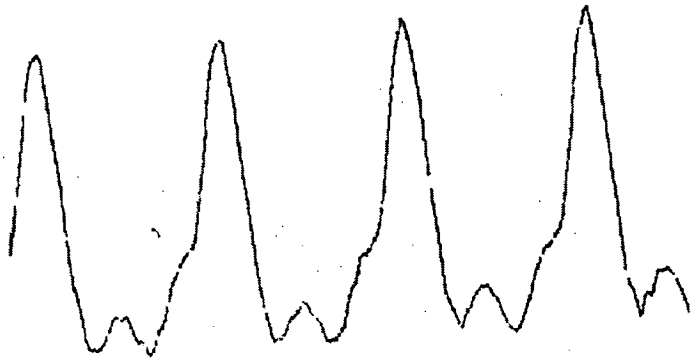
---

[۱] -pharynx
[۲]- coricoid
[۳]- thyroid

واقع شده است، تارآواها در داخل تیروئید و روی حلقه‌ی کری کوئید به صورت دو پرده ماهیچه‌ای بسیار نازک ولی پهن به شکل سرپوش قرار گرفته‌اند. فضای بین تارآواها (تارهای صوتی) را چاک نای می گویند. از دیدگاه زیست شناسی و ظیفه‌ی اصلی حنجره، ایجاد مکانیزم باز و بسته شدن مدخل ششها و در واقع شیر کنترل شش‌هاست. باز و بسته شدن سریع و متوالی تارآواها پدیده واک را ایجاد می کند مکانیزم واک سازی تارهای صوتی بدین گونه است که چنانچه تار آواها به هم نزدیک شوند و راه عبور هوا را مسدود کنند، اگر باز شدن آن مستلزم فشار زیادی نباشد، با اندک فشار هوای شش‌ها تارهای صوتی اندکی باز می شوند و سپس به دلیل کاهش فشار بسته می شوند این عمل بطور سریع و متوالی ادامه می یابد (بیش از ۱۶ بار در هر ثانیه) و به این ترتیب موج صوتی ایجاد می شود. ساختمان و ضخامت تارهای صوتی فرکانس صحبت را تعیین می‌کند. شدت صوت به درجه باز شدگی تارهای صوتی و در نتیجه به مقدار هوای خارج شده بستگی دارد. چنانچه تارهای صوتی محکم بسته شده باشند، باز شدن آنها ممکن است بصورت انفجاری صورت گیرد و چنانچه تارهای صوتی بگونه ای به یکدیگر نزدیک شوند که چاکنای بصورت یک شکاف باریک شود فشار هوا به هنگام خروج ایجاد سایش می کند. این عمل در هنگام نجوا وجود دارد. تولید یکی از آواهای زبان فارسی، با نشانه نوشتاری ((ه )) یا ((ح)) بر اثر این مکانیزم می باشد. واج هایی که در هنگام تولیدشان تارهای صوتی در حالت واکسازی قرار دارند آوایی یا صدادار گویند. برای مثال واج /e/ یک واج آوایی است. شکل سیگنال زمانی آوای /e/ در شکل (۲-۱) مشاهده می‌شود.




شکل ۱-۱ اندام گویایی انسان



شکل ۲-۱ سیگنال زمانی سیگنال /e/

واج هایی که هنگام تولیدشان تارهای صوتی جدا از هم می ایستند و چاکنای باز است و هوا بدون برخورد به حنجره می گذرد را غیرآوایی یا بی صدا گویند، مانند واج (س) که در شکل (۳-۱) مشاهده می‌شود.

شکل ۳-۱ سیگنال زمانی واج /س/

## ٤-۱ اندام های گویایی :

اندام های این بخش از سه حفره حلق ، بینی و دهان تشکیل شده است . از حنجره به بالا تا حفره های بینی گلوگاه نامیده می شود ، که از سه بخش حلق ، ملاز و گلو تشکیل شده است . حفره دهان از کام ، زبان و دندانها تشکیل شده است . حفره های بینی در بالای گلوگاه قرار دارند . دو حفره ذکر شده از جلو به بینی و از عقب به گلوگاه متصل است.

در مرکز تقاطع این سه حفره ، اندام دیگری به نام زبانک[1] قرار دارد که وظیفه آن قطع و وصل ارتباط محفظه ی بینی ( حفره خیشومی [2]) با مجرای صوتی می باشد و در صورت برقراری این اتصال آوای خیشومی تولید خواهد شد .

نکته قابل توجه درباره اندام های گویایی این است که مشخصات اندام های گویایی در تعیین فرکانس صحبت مؤثر هستند .

## ۵-۱ آواشناسی زبان فارسی :

برای نشر صوت هر زبان باید ویژگی های خاص آن زبان را شناخت تا بتوان تکنیک مناسب را برای آن بکار برد و از روش های تفکیک مناسب استفاده کرد . این امر در ترکیب واج ها و تشکیل کلمات و یا قوانین ترکیب کلمات و تشکیل جمله کاربرد مهمتر و آشکارتری دارد . آواهای گوناگون صوتی در هر زبان بر اثر حرکات گوناگون اندام های گفتاری به وجود می آیند. اولین برخورد هوا در لوله های صوتی با تارآواهاست و بسته به وضعیت تارهای صوتی واج های آوایی و غیر آوایی ایجاد می شوند.

### ۱-۵-۱ آوایی و غیرآوایی[3]

چنانچه واجی هنگام ادا شدن همراه با ارتعاش تارهای صوتی باشد آن واج از نوع آوایی خواهد بود و در غیر اینصورت آن واج غیر آوایی خواهد بود. ممکن است واج های آوایی در مجاورت با واج های غیر آوایی قسمتی از واک خود را از دست دهند. همچنین واج های غیر آوایی ممکن است در اثر مجاورت با واج های آوایی، واکدار (جزئی یا کلی) گردند.

### ۲-۵-۱ همخوان ها[4] و واکه ها[5]

آواهای زبان فارسی را می توان به دو دسته همخوان و واکه تقسیم کرد. همخوان ها، آواهایی هستند که در آغاز کلمه واقع می‌شوند و هنگام تولید آنها در نقطه ای از دستگاه صوتی گرفتگی یا تنگی که منجر به سایش گردد به وجود می آید.

---

۱-volume
۲-Nasal Cavity
۳- Voiced / Unvoiced
٤-Consonant
٥-Vowel

همخوانهای زبان فارسی به وسیله جریان بازدم تولید می شوند. ممکن است در این حالت راه عبور هوا به کلی بسته یا باز باشد. در حالت مکانیزم باز می توانند خیشومی[1] یا دهانی باشند. و در مکانیزم بسته که عموما دهانی هستند ممکن است انفجاری یا انفجاری همراه با سایش یا لرزش باشند. در فارسی ۲۳ همخوان وجود دارند که عبارتند از:

B,p,t,d,s,z,č,g,q,?,v,š,x,ž,f,h,m,n,I,y

جدول (۱-۱) کلیه واج گونه های همخوانی زبان فارسی را نشان می دهد .

جدول ۱-۱ واجگونه های همخوانی زبان فارسی[۱]

| واجگاه | وجود واک | معادل نوشتاری | واج ها |
|---|---|---|---|
| دولبی | بیواک | پ | /p/ |
| دولبی | واکدار | ب | /b/ |
| دندانی | بیواک | ت،ط | /t/ |
| دندانی | واکدار | د | /d/ |
| نرمکامی | بیواک | ک | /k/ |
| نرمکامی | واکدار | گ | /g/ |
| ملازی | واکدار | ق،غ | /q/ |
| چاکنایی | بیواک | ع ، ء | /ʔ/ |
| لیوی | بیواک | ث،س،ص | /s/ |
| لیوی | واکدار | ذ،ز،ظ،ض | /z/ |
| کامی- لیوی | بیواک | ش | /š/ |
| کامی-لیوی | واکدار | ژ | /ž/ |
| لب و دندانی | بیواک | ف | /f/ |
| لب و دندانی | واکدار | و | /v/ |
| ملازی | بیواک | خ | /x/ |
| چاکنایی | بیواک | ح ، ه | /h/ |
| کامی-لیوی | بیواک | چ | /č/ |
| کامی- لیوی | واکدار | ج | /j/ |
| لیوی | واکدار | ر | /r/ |
| دولبی | واکدار | م | /m/ |
| لیوی | واکدار | ن | /n/ |
| لیوی(کناری) | واکدار | ل | /l/ |
| سخت کامی | واکدار | ی | /y/ |

۱- Nasal

واکه یک آوای مداوم صدادار است که جریان هوا بدون برخورد به مانع در مجرای صوتی عبور می کند. عموما واکه ها همگی دهانی هستند، هوای به ارتعاش درآمده بوسیله تارهای صوتی در ایجاد واکه وجود دارد و این امر سبب ارتعاش هوای دهان می شود، بنابراین شکل های گوناگون حفره های دهان سبب ایجاد کیفیت های مختلف آوایی و واکه های مختلف می شود. در زبان فارسی ۸ واکه داریم که ٦ واکه بسیط و دو واکه مرکب هستند. هنگام ادای یک واکه ی مرکب اندام های گفتاری معمولا از موقعیت یک واکه در جهت واکه ی دوم حرکت می کنند و به طور کامل در وضعیت تولیدی واکه ی دوم قرار نمی گیرند. جدول (۲-۱) انواع واج گونه های زبان فارسی (واکه های بسیط) را نشان می دهد.

جدول ۲-۱ واج‌گونه های واکه‌ای زبان فارسی[۱]

| واجگونه ها | واج ها |
|---|---|
| ی ، ای | /i/ |
| کسره | /e/ |
| فتحه | /a/ |
| او ، و | /u/ |
| ضمه | /o/ |
| آ ، ا ، وا | /â/ |

## ۳-۵-۱ واج[۱] و واج گونه[۲]:

هر یک از همخوان ها و یا واکه ها شامل گروهی از آواهاست که با وجود شباهت اختلافاتی نیز با یکدیگر دارند که به آنها واج گونه گویند و به هر یک از همخوان و یا واکه ها واج گفته می شود . برای مثال [c] (پیشین یا گسترده) در واژه [cine] (کین) عضو گروه [k] با نشانه [k] می باشد . [c] را یک واج گونه و [k] را یک واج گویند[۱].

قبل از به پایان رساندن این بحث چند اصطلاح مهم که در حوزه ی آواشناسی مطرح است و در پردازش دیجیتال سیگنال صوت از آنها استفاده خواهیم کرد توضیح داده می شود.

۱- فرکانس های فرمنت[۳]:

۱-Phoneme
۲-Allophone
۳-Formant

مهمترین تفاوت بین فونم ها را می توان در طیف فرکانسی آنها مشاهده کرد ،مجرای صوتی مانند یک حفره رزونانس با فرکانس های طبیعی عمل می کند که فرکانس های آن با توجه به محل تولید صدا در مجرای صوتی تغییر می کند.این فرکانس های رزونانس، فرمنت نامیده می شوند که به صورت مکان هایی با انرژی زیاد در اسپکتوگرام مشاهده می شودو یا پیک هایی که در طیف فرکانسی دیده می شود فرمنت مربوط به مجرایی صوتی می باشد.

۲- گام[۱]:

اگر تار آواها قدری کشیده شوند ولی فشار هوای درون شش ها بیش از نیرویی باشد که تار آواها را به هم چسبانده است پدیده ای به نام تولید واک رخ می دهد که اهمیت زیادی در آواشناسی دارد،وجود واک تابع زمانی سیگنال صحبت رابه شکل تناوبی در می آورد که به دوره تناوب این سیگنال گام با فرکانس اصلی(f۰) نیز گفته می‌شود.

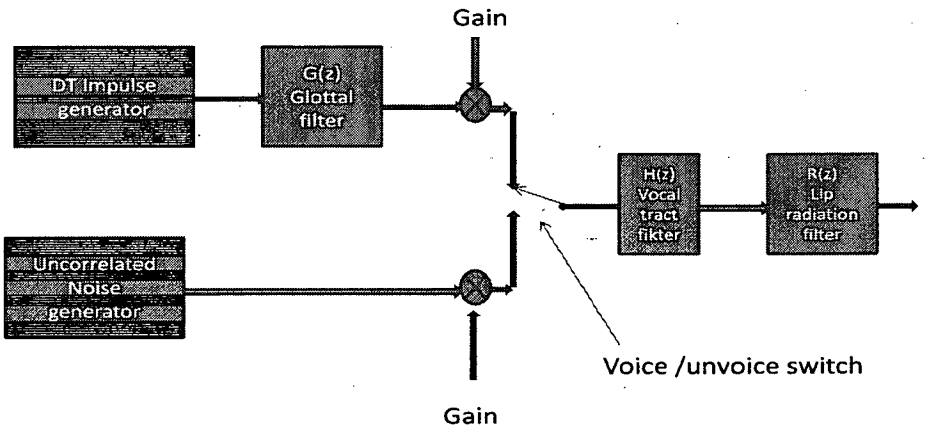
در زبان فارسی حدود $\frac{۲}{۳}$ واج ها واکدار و $\frac{۱}{۳}$ بقیه بدون واک هستند.

## ٦-۱ مدل نظری سیستم گفتار انسان:

همانطور که در بخش های قبل ذکر شد سیستم طبیعی گفتار انسان از یک مجرای صوتی تشکیل شده که از تارها تا لبها امتداد دارد و ورودی این سیستم نیز در اثر عبور هوا از ریه به دهان تامین می شود . سطح مقطع لوله صوتی انسان در هنگام صحبت تغییر می کند و ثابت نمی باشد . در این صورت می توان لوله صوتی را که تقریبا مستقل از ورودی آن است در بلوک جداگانه مدلسازی کرد . شکل٤-۱ مدل سیستم صوتی انسان را نشان می دهد .

---

۱-Pitch




شکل ٤-١ مدل نظری سیستم گویایی انسان

مقادیر G (gain) و ضرایب فیلتر تمام قطب دیجیتال با توجه به مدل انتخابی برای دستگاه صوتی انتخاب می‌شوند. چون مدل انتخابی ، همانطور که در فصل ۳ توضیح داده خواهد شد مدل فیلتر پیشگویی خطی است این پارامترها را این مدل پیش بینی می کند و منبع تولید صدا که تقریبا مستقل از شکل اندام گویایی تصور می شود بسته به اینکه آوای ادا شده از نوع صدادار (voiced) و یا از نوع بی صدا (unvoiced) باشد ، به ترتیب مولد قطار پالس یا مولد نویز سفید انتخاب می شود . گرچه پارامترهای فیلتر تمام قطب دیجیتال متغیر با زبان است اما ثابت شده است که می توان آنها را در محدوده های ۱۰ الی ۲۵ میلی ثانیه تقریبا ثابت تصور کرد[٤][۱۱] . در این صورت در یک فریم با این طول زمانی می توان این پارامترها را ثابت فرض کرد . توضیحات بیشتر در باره ی مدل سیستم صوتی انسا ن وچگونگی استفاده از این مدل در پردازش سیگنال صوت در فصل های بعدی ارایه خواهد شد.

## ۱-۷ خلاصه:

در این فصل ابتدا اندام های صوتی انسان به صورت اجمال معرفی شد ونقش هریک در چگونگی تولید سیگنال صوت بیان شد.در ادامه تقسیم بندی سیگنال های صوت به واج های آوایی وغیر آوایی توضیح داده شد و واج های آوایی وغیر آوایی زبان فارسی در دو جدول جداگانه آورده شد . اصطلاحات مطرح در علم آواشناسی در حد نیاز معرفی شد.(اصطلاحات مورد نیاز دیگر در فصول بعدی و در هنگام نیاز توضیح داده خواهد شد)و در پایان مدل استفاده شده برای سیستم صوتی انسان که در پردازش سیگنال صوت کارایی دارد ارایه گردید.