

آواشناسی آکوستیک و شنیداری

نوشتۀ: کیث جانسون

ترجمۀ:

دکتر زهرا محمودزاده

دکتر رحیمه روح پرور



پژوهشگاه علوم و فناوری اطلاعات ایران

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

آواشناسی

آکوستیک و شنیداری

نوشته:

کیث جانسون

ترجمه:

زهرا محمودزاده

هیئت علمی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

رحیمه روح پرور

هیئت علمی دانشگاه شهید باهنر کرمان

ویراستار:

حسین مختاری معمار



وزارت علوم، تحقیقات و فناوری
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

سرشناسه	: جانسون، کیت، ۱۹۴۴ - م. Johnson, Keith
عنوان و نام پدیدآور	: آواشناسی آکوستیک و شنیداری/نوشته کیت جانسون؛ ترجمه زهرا محمودزاده، رحیمه روح‌پرور؛ ویراستار حسین مختاری‌معمار.
مشخصات نشر	: تهران: پژوهشگاه علوم و فناوری اطلاعات ایران: چاپار، ۱۳۹۴.
مشخصات ظاهری	: ۳۴۶ص.
شابک	: 978-964-7519-93-9
وضعیت فهرست نویسی	: فیبا
یادداشت	: عنوان اصلی: Acoustic and auditory phonetics, 2013.
موضوع	: آواشناسی فیزیکی
موضوع	: شنوایی
شناسه افزوده	: محمودزاده، زهرا، ۱۳۵۴ - مترجم
شناسه افزوده	: روح‌پرور، رحیمه، ۱۳۵۷ - مترجم
شناسه افزوده	: مختاری‌معمار، حسین، ۱۳۲۶ - ویراستار
شناسه افزوده	: پژوهشگاه علوم و فناوری اطلاعات ایران
رده بندی کنگره	: ۱۳۹۴ ج۲۸/۵/۲۲۱p
رده بندی دیویی	: ۴۱۴
شماره کتابشناسی ملی	: ۴۰۲۸۴۸۵

این کتاب از سلسله انتشارات پژوهشگاه علوم و فناوری اطلاعات ایران است که با همکاری انتشارات چاپار منتشر شده است.

آواشناسی آکوستیک و شنیداری



نوشته: کیت جانسون

ترجمه: زهرا محمودزاده، رحیمه روح‌پرور

ویراستار: حسین مختاری‌معمار

© ۱۳۹۴: پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

صفحه آرا و طراح جلد: رحیم کیبیرصابر

نوبت چاپ: اول ۱۳۹۴ | شمارگان: ۱۰۰۰ نسخه

تهران - خیابان ولی عصر (عج) - بالاتر از زرتشت

کوچه نوربخش - پلاک ۴۰ - نشر چاپار

تلفن پخش: ۸۸۸۹۹۶۸۰ - ۰۹۱۲۱۹۶۰۲۱۴

بها: ۲۰۰.۰۰۰ ریال

فهرست مطالب

پیشگفتار	۹
مقدمه	۱۳
بخش ۱: مبانی	۱۷
فصل ۱: مبانی آکوستیک و صافی‌های آکوستیکی	۱۹
۱-۱. احساس صدا	۱۹
۲-۱. انتشار صوت	۲۰
۳-۱. اقسام صوت	۲۴
۱-۳-۱. امواج متناوب ساده	۲۴
۲-۳-۱. امواج متناوب مرکب	۲۷
۳-۳-۱. امواج نامتناوب	۳۳
۴-۱. صافی‌های آکوستیکی	۳۶
منابعی برای مطالعه بیشتر	۴۰
تمرین	۴۱
فصل ۲: نظریه آکوستیکی تولید گفتار: تولید شوا	۴۳
۱-۲. واک	۴۳
۲-۲. مقادیر ذره‌ای واک	۴۷
۳-۲. صافی دستگاه گفتار	۵۰
۴-۲. پاندول، امواج ایستاده و سازه‌های واکه	۵۲
۵-۲. جستجوی گره‌ها و پادگره‌ها در یک لوله آکوستیکی	۷۰
منابعی برای مطالعه بیشتر	۷۳
تمرین	۷۴
فصل ۳: پردازش دیجیتالی سیگنال	۷۵
۱-۳. سیگنال‌های پیوسته در مقابل سیگنال‌های گسسته	۷۵
۲-۳. تبدیل آنالوگ به دیجیتال	۷۷
۱-۲-۳. نمونه‌برداری	۷۸
۲-۲-۳. کمی‌سازی	۸۳
۳-۳. روش‌های تجزیه سیگنال	۸۹

۹۰.....	۳-۱. دامنه آرام‌اس
۹۲.....	۳-۲. تبدیل سریع فوریه
۹۷.....	۳-۳. محاسبه زیربمی با روش خودهمبستگی
۱۰۲.....	۳-۴. صافی‌های دیجیتال
۱۰۷.....	۳-۵. کدگذاری پیش‌گویانه خطی (ال‌پی‌سی)
۱۱۶.....	۳-۶. طیف و طیف‌نگاشت
۱۲۰.....	منابعی برای مطالعه بیشتر
۱۲۱.....	تمرین
۱۲۳.....	فصل ۴: مبانی شنوایی
۱۲۴.....	۴-۱. آناتومی سیستم شنوایی بیرونی
۱۲۵.....	۴-۲. احساس شنیداری بلندی صوت
۱۳۱.....	۴-۳. پاسخ فرکانسی سیستم شنوایی
۱۳۴.....	۴-۴. اشباع و استتار
۱۳۸.....	۴-۵. بازنمایی‌های شنیداری
۱۴۳.....	منابعی برای مطالعه بیشتر
۱۴۵.....	تمرین
۱۴۷.....	فصل ۵: درک گفتار
۱۴۹.....	۵-۱. توانایی شنیداری درک گفتار را شکل می‌دهد
۱۵۴.....	۵-۲. دانش آوایی درک گفتار را شکل می‌دهد
۱۵۴.....	۵-۲-۱. درک مقوله‌ای
۱۶۰.....	۵-۲-۲. انسجام آوایی
۱۶۵.....	۵-۳. دانش زبانی درک گفتار را شکل می‌دهد
۱۷۰.....	۵-۴. شباهت درکی
۱۷۲.....	۵-۴-۱. رسم نقشه بر اساس فاصله
۱۷۴.....	۵-۴-۲. نقشه درکی همخوان‌های سایشی
۱۸۲.....	منابعی برای مطالعه بیشتر
۱۸۵.....	تمرین
۱۸۷.....	بخش ۲: تجزیه گفتار
۱۸۹.....	فصل ۶: واکه‌ها
۱۸۹.....	۶-۱. مدل‌های لوله‌ای تولید واکه
۱۹۷.....	۶-۲. نظریه اختلال
۲۰۳.....	۶-۳. واکه‌های مرجع، نظریه ذره‌ای و نظریه پراکندگی تطبیقی
۲۰۵.....	۶-۴. سازه‌ها و فضای واکه‌ای آکوستیکی
۲۰۹.....	۶-۵. بازنمایی شنیداری و آکوستیکی واکه‌ها
۲۱۰.....	۶-۶. درک زبان‌گذر واکه
۲۱۴.....	منابعی برای مطالعه بیشتر
۲۱۶.....	تمرین

فصل ۷: سایشی‌ها.....	۲۱۹
۱-۷. اغتشاش.....	۲۱۹
۲-۷. محل تولید سایشی‌ها.....	۲۲۶
۳-۷. نظریه ذره‌ای و سایشی‌ها.....	۲۳۱
۴-۷. طیف‌های شنیداری سایشی.....	۲۳۴
۵-۷. ابعاد درکی سایشی‌ها.....	۲۳۸
منابعی برای مطالعه بیشتر.....	۲۴۱
تمرین.....	۲۴۲
فصل ۸: انسدادی‌ها و انسایشی‌ها.....	۲۴۳
۱-۸. توابع منبع در انسدادی‌ها و انسایشی‌ها.....	۲۴۴
۱-۱-۸. انواع واک‌سازی.....	۲۴۴
۲-۱-۸. منبع صوت در انسدادی‌ها و انسایشی‌ها.....	۲۴۸
۲-۸. توابع صافی دستگاه گفتار در انسدادی‌ها.....	۲۵۳
۳-۸. انسایشی‌ها.....	۲۵۷
۴-۸. ویژگی‌های شنیداری انسدادی‌ها.....	۲۵۸
۵-۸. درک انسدادی‌ها در بافت واکه‌ای.....	۲۶۰
منابعی برای مطالعه بیشتر.....	۲۶۳
تمرین.....	۲۶۴
فصل ۹: خیشومی‌ها و کناری‌ها.....	۲۶۷
۱-۹. پهنای باند.....	۲۶۷
۲-۹. انسدادی‌های خیشومی.....	۲۷۰
۳-۹. همخوان‌های کناری.....	۲۸۱
۴-۹. خیشومی‌شدگی.....	۲۸۴
۵-۹. درک همخوان خیشومی.....	۲۸۹
منابعی برای مطالعه بیشتر.....	۲۹۲
تمرین.....	۲۹۴
منابع.....	۲۹۵
پاسخ به منتخبی از سؤالات.....	۳۰۳
واژه‌نامه فارسی به انگلیسی.....	۳۱۱
واژه‌نامه انگلیسی به فارسی.....	۳۲۵
نمایه.....	۳۴۱

پیشگفتار

کتاب «آواشناسی آکوستیک و شنیداری» ترجمه اثر تخصصی و ارزشمند «کیث جانسون» استاد زبانشناسی دانشگاه برکلی در آمریکاست. پیتز لدفولگد در توصیف این کتاب که بر روی جلد آن آمده، می‌نویسد: «معلوم شد این کتاب درسی با طرح‌ریزی خوب و استادانه‌اش برای دانشجویانی که می‌خواهند بیشتر از مبانی آواشناسی آکوستیک بدانند، بسیار ارزشمند و مفید است.» و جنت پیرهامبرت، استاد دانشگاه نورث‌وسترن در آمریکا می‌افزاید: «بهترین کتاب درسی آواشناسی است که تاکنون به کار برده‌ام، به‌طور بی‌نظیری موفق در نشان دادن شیوه استدلال در مدل‌سازی آکوستیکی به مبتدیان».

کتاب شامل دو بخش است. بخش اول که به مبانی آواشناسی آکوستیک اختصاص یافته، شامل پنج فصل است. در فصل اول، مبانی فیزیک امواج صوتی، شکل موج آنها و انواع فیلترهای آکوستیکی توضیح داده می‌شود. در فصل دوم، با پرداختن به مبانی فیزیک بازخوان‌ها یا سازه‌های دستگاه گفتار، روش و فرمول محاسبه فرکانس‌های بازخوانی دستگاه گفتار توضیح داده می‌شود. در فصل سوم، با پرداختن به مبانی فنی سیگنال دیجیتال، به توصیف ریاضی تکنیک‌ها یا روش‌هایی در پردازش دیجیتالی سیگنال می‌پردازد که در آواشناسی آکوستیک نیز برای تجزیه و تحلیل سیگنال آکوستیکی به کار می‌روند، روش‌هایی مثل تجزیه فوریه، ال‌پی‌سی، طیف‌نگاشت و غیره. فصل چهارم، در مبحث مربوط به مبانی شنوایی، تفاوت‌های

شنیداری سیگنال صوتی با سیگنال آکوستیکی مطرح می‌شود و تأثیر سیستم شنوایی بر فرکانس، دامنه و طیف امواج صوتی و نیز بر واحدهای آوایی توضیح داده می‌شود. فصل پنجم، شواهدی از نقش سیستم شنیداری، دانش آوایی و دانش زبانی در درک گفتار ارائه می‌شود و با توصیف نقشه درکی، به معرفی روشی برای محاسبه میزان شباهت درکی بین آواها به عنوان عاملی کلیدی در تحول آوایی زبان و تناوبات واجی می‌پردازد.

بخش دوم که به تجزیه و تحلیل گفتار اختصاص دارد، شامل چهار فصل واکه‌ها، سایشی‌ها، انسدادی‌ها و انسایشی‌ها و بالاخره خیشومی‌ها و کناری‌ها است. در همه این فصل‌ها، مدل آکوستیکی به‌ویژه مدل لوله‌ای تولید آواها براساس نظریه منبع-صافی توضیح و طیف شنیداری و نقشه درکی طبقات آوایی نشان داده می‌شود. علاوه بر آن، در فصل‌های مربوط به واکه‌ها و سایشی‌ها، به توضیح ویژگی‌های نگاشت آکوستیکی تولید واکه‌ها و سایشی‌ها براساس نظریه ذره‌ای استیونس (۱۹۷۲) می‌پردازد. در فصل مربوط به خیشومی‌ها و کناری‌ها نیز تولید پادسازه‌ها در دستگاه گفتار حین تولید این طبقات آوایی توضیح و نشان داده می‌شود.

نویسنده کتاب «آواشناسی آکوستیک و شنیداری» از زاویه‌های مختلف فیزیک، ریاضی، فنی و مهندسی، شنوایی، درک و غیره به موج صوتی گفتار می‌پردازد. از این رو لاجرم بسیاری از موضوعات و اصطلاحات تخصصی و فنی کتاب برگرفته از سایر شاخه‌های تخصصی علمی به جز آواشناسی و زبان‌شناسی است. معادل فارسی بسیاری از اصطلاحات آواشناسی با مراجعه به منابع موجود آواشناسی به‌ویژه آثاری مانند «آواشناسی زبان فارسی» (یدالله ثمره، ۱۳۸۵)، «آواشناسی» (علی محمد حق‌شناس، ۱۳۸۶) و «نظام آوایی زبان فارسی» (محمود بی‌جن‌خان، ۱۳۹۲) بدست آمده است. معادل واژه‌ها و اصطلاحات تخصصی سایر شاخه‌های علمی از منابع متعدد به‌ویژه با مراجعه به تزاروس برخط پژوهشگاه علوم و فناوری اطلاعات ایران بدست آمده است. این تزاروس حاوی واژگان و اصطلاحات حوزه‌های مختلف

علمی مانند فنی و مهندسی و علوم پایه است. علاوه بر آن برخی معادل‌ها نیز با مراجعه به برخی فرهنگ‌های تخصصی مانند فرهنگ جامع کامپیوتر و فناوری اطلاعات (ایرج صفا، ۱۳۸۸)، فرهنگ تشریحی کامپیوتر، شبکه، تکنولوژی اطلاعات، اینترنت، مخابرات، بی‌سیم، فیبر نوری، ماهواره (ترجمه محمدحسن مهدوی، ۱۳۸۸)، و فرهنگ توصیفی آسیب‌شناسی گفتار و زبان (رضا نیلی‌پور و همکاران، ۱۳۹۲) بدست آمده است. معادل شماری از اصطلاحات نیز برگرفته از واژه‌های مصوب فرهنگستان است که در وب‌سایت آن در دسترس است. در موارد معدودی نیز معادل فارسی کلمه پیشنهاد مترجم است.

معادل انگلیسی واژه‌ها و اصطلاحات تخصصی در پاورقی‌ها آمده است. با فرض این که خواننده آشنایی مقدماتی با آواشناسی آکوستیک دارد، به ذکر معادل انگلیسی برخی واژه‌های آواشناسی آکوستیک فقط در انتهای کتاب و در بخش واژه‌نامه اکتفا شده و از درج آنها در پاورقی خودداری شده است. درواقع هدف از درج معادل انگلیسی واژه‌ها در پاورقی، کمک به خواننده برای درک و فهم بهتر و بیشتر متن و رفع ابهام است. ازاین‌رو حذف معادل انگلیسی برخی از واژه‌های تخصصی دیگر در پاورقی‌ها با این دیدگاه صورت گرفته که معادل فارسی این واژه‌ها خود رسا و گویاست؛ به عنوان مثال، اصطلاحاتی مثل «موج انعکاسی» یا «موج انتشار». بنابراین، واژه‌نامه انتهای کتاب حاوی اصطلاحات بیشتری است نسبت به آنچه در پاورقی‌ها دیده می‌شود.

در برخی موارد سعی شده با آوردن مختصر توضیحی درباره واژه‌ها یا مباحث تخصصی در پاورقی، اطلاعاتی هرچند اندک در اختیار خواننده قرار گیرد. توضیحاتی که به‌زعم مترجم به‌ویژه برای روشن‌تر شدن معنا و مفهوم واژه‌ها یا مباحث تخصصی سایر حوزه‌های علمی به جز آواشناسی و زبان‌شناسی سودمند به نظر می‌رسید. موارد معدودی نیز در متن اضافه شده که بین دو قلاب قرار گرفته است. در هر رو تلاش شد با حفظ اصل امانت‌داری در ترجمه، مفاهیم مورد نظر نویسنده منتقل شود.

کتاب حاضر می‌تواند به عنوان یک منبع درسی مورد استفاده پژوهشگران، اساتید و دانشجویان رشته‌های زبان‌شناسی همگانی، زبان‌شناسی رایانشی، آسیب‌شناسی گفتار و زبان، گفتاردرمانی، شنوایی‌سنجی و شاخه‌هایی از هوش مصنوعی و مهندسی برق که با زبان و گفتار سروکار دارند، قرار گیرد.

ترجمه این اثر بدون کمک و یاری مهندس جلال معصومی که در فهم بسیاری از مباحث تخصصی فنی و ریاضی همراهی نمود و نیز استاد ارجمند دکتر بری هسلوود (Barry Hesselwood)، استاد بازنشسته آواشناسی در دانشگاه لیدز انگلستان که هر جا مفهوم دچار پیچیدگی زبان می‌شد، راه‌گشا بود، میسر نبود. از جانب آقای معماری مختار، ویراستار محترم کتاب، خانم نادیا حاجی عزیزی و خانم مینا سادات قدیری، مدیر و کارشناس دفتر ارتباطات علمی و نیز سایر دست‌اندرکاران پژوهشگاه علوم و فناوری اطلاعات ایران که شرایط ترجمه و چاپ کتاب حاضر را فراهم آوردند، نهایت تشکر و قدردانی را داریم.

تابستان ۱۳۹۴

محمودزاده - روح‌پرور

مقدمه

این کتاب مقدمه‌ای کوتاه و غیرفنی (مناسب به عنوان یک منبع مکمل برای درس آواشناسی عمومی یا علم گفتار) در چهار مبحث مهم آواشناسی آکوستیک شامل (۱) ویژگی‌های آکوستیکی طبقات عمده آواهای گفتاری، (۲) نظریه آکوستیکی تولید گفتار، (۳) بازنمایی شنیداری گفتار و (۴) درک گفتار است. من این کتاب را برای دانشجویان در درس‌های مقدماتی آواشناسی زبان‌شناسی، علوم گفتاری و شنیداری و شاخه‌هایی از مهندسی برق و روان‌شناسی شناختی که با گفتار سروکار دارند، نوشتم. پنج فصل اول به معرفی مبانی آکوستیک، نظریه آکوستیکی تولید گفتار، پردازش دیجیتالی سیگنال، شنوایی و درک گفتار می‌پردازد و در چهار فصل باقیمانده، طبقات عمده آوایی مورد مطالعه قرار می‌گیرد و مشخصه‌های آکوستیکی آواهای گفتاری طبق پیش‌بینی نظریه آکوستیکی تولید گفتار، ویژگی‌های شنیداری و مشخصه‌های درکی آنها مرور می‌شود. در پایان هر فصل، فهرستی از منابع پیشنهادی برای مطالعه بیشتر به همراه تعدادی تمرین آمده است. تمرین‌ها با تأکید بر اصطلاحاتی که در هر فصل با حروف سیاه مشخص گردیده (و در قسمت «واژگان تخصصی» فهرست شده‌اند)، خواننده را به به‌کاربردن مفاهیم مطرح‌شده در فصل ترغیب می‌نمایند. بعضی از تمرین‌ها در اصل جنبه مرور مطالب را دارند، اما بسیاری از آنها به مسائل یا موضوعاتی می‌پردازند که به‌طور صریح در متن به آنها اشاره نشده است. پاسخ به برخی از این تمرین‌ها در انتهای کتاب آمده است.

در متن، پیام‌های پنهانی نیز گنجانده‌ام؛ از آن جمله: (۱) نمونه‌های آوایی از زبان‌ها و گویشوران متنوعی گرفته شده‌اند، زیرا برون‌داد آکوستیکی دستگاه گفتار فقط به اندازه و شکل دستگاه گفتار و به مکانیسم‌های آیرودینامیکی تولیدکنندهٔ نوفه بستگی دارد. این جنبه‌های گفتاری به‌وسیلهٔ آناتومی و فیزیولوژی شکل می‌گیرند، بنابراین فراسوی عادات فرهنگی یا شخصی قرار دارند. (۲) این کتاب، کتابی است دربارهٔ آواشناسی آکوستیک و شنیداری، زیرا تحلیل آکوستیکی استاندارد، حقایق زبانی محدود و ناقصی را نشان می‌دهد. سیستم شنوایی، سیگنال گفتاری را به‌شیوه‌های بسیار جالبی تغییر می‌دهد و اگر بخواهیم پی به ارزش زبانی آکوستیک گفتار (یا بی‌ارزشی آن) ببریم، لازم است به سیستم شنوایی نیز توجه کنیم. ارزش زبانی آواشناسی آکوستیک از سویی تحت تأثیر پردازش درکی شناختی نیز قرار دارد، از این رو هر یک از فصل‌های نیمهٔ دوم کتاب، جنبه‌ای از درک گفتار را مورد توجه قرار می‌دهد. (۳) در کتاب، فرمول‌هایی دیده می‌شود. در واقع، برای حل برخی تمرین‌ها در انتهای فصل‌ها نیاز به استفاده از ماشین حساب است. استفاده از فرمول ممکن است نوعی شانه‌خالی کردن از کار تلقی شود، اما زبان ریاضی اغلب بسیار دقیق‌تر از هر نثری است که بتوان تصور نمود. در دفاع از خود، باید بگوییم که فقط دو فرمول اصلی در کتاب ملاحظه می‌شود (برای بازخوان لوله‌هایی که یا در هر دو انتها یا فقط در یک انتها بسته هستند)، علاوه بر آن، بخش واقعاً جالب آواشناسی آکوستیک وقتی شروع می‌شود که ماشین حساب خود را بیرون می‌آورید. ریاضیات این کتاب (هرچند اندک) ساده است. (۴) در سرتاسر کتاب از علائم آی‌پی‌ای (انجمن بین‌المللی آواشناسی) استفاده شده است. فرض بر این بوده که خواننده دست‌کم آشنایی اجمالی با مجموعهٔ متداول علائم آوانویسی دارد.

مطالب نیمه مرتبط در مستطیل‌ها

مباحث متنوع جالبی در حاشیه موضوعات اصلی فصل‌ها قابل طرح است. از این رو، گاهی کتاب در مستطیل‌هایی سایه‌دار مثل همین نمونه، از موضوع اصلی فاصله گرفته تا به گزیده‌ای از سؤالاتی که دانشجویان ممکن است پرسیده باشند (مهم‌ترین سؤالات)، به شکلی ساده پاسخ دهد. مطالب این مستطیل‌ها، از گفتار در زیر آب گرفته تا درک پادسازه‌ها، تا رقم‌های دیجیتالی و آیرودینامیک بزرگراه‌ها را در بر می‌گیرد. من این مطالب جانبی را در کتاب گنجانده‌ام، زیرا هیچ سؤالی آنقدر ساده نیست که مطرح نشود. برخی از جالب‌ترین مطالب کتاب را ممکن است در این مستطیل‌ها ببینید.

اصلاحاتی در چاپ سوم

از خوانندگان، استادان و دانشجویانی که با اظهارنظرهای خود امکان بهترکردن کتاب را فراهم آورده‌اند، تشکر می‌کنم. تغییرات عمده‌ای که استادان ملاحظه خواهند نمود عبارت‌اند از: (۱) تغییر در ترتیب فصل‌ها، معرفی نظریه آکوستیکی تولید گفتار در کتاب جلوتر آمده و اشاره به شنوایی و درک گفتار نیز زودتر مطرح گردیده است. متوجه شدم که دلیل خوبی برای گذاشتن فصل‌های شنوایی و درک گفتار در بخش‌های انتهایی کتاب وجود دارد و این ترتیبی که انتخاب کرده‌ام، پیچیدگی‌ها و دشواری‌هایی را برای استادان پیش می‌آورد. امیدوارم که نتیجه این تغییر، یعنی امکان جمع‌آوری داده‌های آکوستیکی، شنیداری و درکی آواهای گفتاری با هم در هر یک از فصل‌های ۶ تا ۹، بتواند آن را جبران کند. (۲) فصل پردازش دیجیتالی سیگنال، به‌روز شده است تا با نرم‌افزارها و سخت‌افزارهایی که اکنون در دسترس است، تناسب بیشتری داشته باشد. بخش تجزیه ال‌پی‌سی نیز بازنویسی شده است. (۳) فصل درک گفتار نیز تغییر کرده است. در این فصل، به مطالب نظری و همین‌طور نکات عملی پرداخته شده است که بخش عمده چاپ دوم را نیز تشکیل می‌داد. در این فصل، دیدگاه خاصی را مطرح کرده‌ام که ممکن است برخی استادان با آن موافق

نباشند، اما در عین حال سعی کرده‌ام فضا را برای ورود استادان به بحث نظری با کتاب (و با دانشجویان) در این موضوع باز بگذارم. (۴) بخش‌هایی از فصل‌هایی که به معرفی نظریه آکوستیکی تولید گفتار و آکوستیک واکه‌ها می‌پردازد، بازنویسی شده تا بازنمایی دقیق‌تر (و صحیح‌تری) از بازخوان‌ها و امواج ایستاده در دستگاه گفتار به دست دهد. (۵) فصل شنوایی شامل یک بخش جدید در مورد اشباع و استتار است. (۶) بسیاری از طیف‌نگاشت‌های کتاب جایگزین طیف‌نگاشت‌هایی شدند که تفسیر آنها در مقایسه با چاپ‌های پیشین کتاب ساده‌تر است. (۷) اکنون هر فصل با گزیده‌ای از منابع پیشنهادی برای مطالعه پایان می‌یابد.

دانشجویان متوجه تغییری بین چاپ سوم با دوم نخواهند شد؛ مگر اینکه شما آدم خیلی عجیب و غریبی باشید و به چاپ‌های قدیمی کتاب‌ها سر بزنید یا اینکه خیلی بدشانس باشید و مجبور شوید درس را بعد از انتشار این چاپ دوباره بگیرید. مثل همیشه، آرزوی من برای دانشجویانی که از این کتاب استفاده می‌کنند این است که آموختن آواشناسی آکوستیک با این کتاب سرگرم‌کننده‌تر و جذاب‌تر باشد تا بدون آن.

کیث جانسون

بخش ۱

مبانی

فصل ۱

مبانی آکوستیک و صافی‌های آکوستیکی

۱-۱. احساس صدا

رخدادهای بسیاری در جهان احساس صدا را به وجود می‌آورند. به عنوان مثال، به هم خوردن درها، نواختن ویلن، زوزه باد و گفتار انسان. همه این صداها و هر صدای دیگری که بتوان تصور کرد، مستلزم نوعی جابه‌جایی است. این جابه‌جایی‌ها سبب تولید نوسانات فشار در هوای اطراف (یا هر رسانای آکوستیکی دیگری) می‌شود. هنگامی که نوسانات فشار به پرده گوش می‌رسد، آن را به لرزش در می‌آورد و سیستم شنیداری گوش انسان، این لرزش‌ها را به تکانه یا پالس‌های عصبی تبدیل می‌کند که ما آنها را به صورت صدا تجربه می‌کنیم. بنابراین، صدا هنگامی تولید می‌شود که نوسانات فشار با پرده گوش برخورد می‌کنند. شکل موج آکوستیکی ثبت نوسانات فشار مولد صدا در واحد زمان است. (برای مطالعه بیشتر در زمینه مطالب این فصل رجوع شود به: لدفوگد^۱، فرای^۲ ۱۹۷۹ و استیونس^۳ ۱۹۹۹)

1. Ladefoged

2. Fry

3. Stevens

رسانای آکوستیکی

معمولاً نوسانات فشاری که به عنوان صدا شنیده می‌شوند، در هوا تولید می‌شوند، اما صدا می‌تواند در سایر رساناهای آکوستیکی نیز منتقل شود. به عنوان مثال، وقتی زیر آب شنا می‌کنید، ممکن است صداهای خفیفی از افراد بالای آب و یا صدای غلغل حباب‌های زیر آب را بشنوید. به جز هوا سایر گازها نیز می‌توانند نوسانات فشاری را انتقال دهند که تولید صدا می‌کند. مثلاً وقتی پس از استنشاق گاز هلیوم درون یک بالون شروع به حرف زدن می‌کنید، صدای شما از طریق گاز هلیوم انتقال پیدا می‌کند که در این حالت با صدای عادی تفاوت دارد. این‌گونه مثال‌ها نشان می‌دهند که ویژگی‌های صدا تا حدی به رسانای آکوستیکی، سرعت انتقال نوسانات فشار در درون رسانه و نیز به میزان مقاومت رسانه در برابر چنین نوساناتی بستگی دارد.

۱-۲. انتشار صوت

نوسانات فشاری که به پرده گوش می‌رسند، احساس صدا را تولید می‌کنند. صدا می‌تواند مسافت‌های نسبتاً طولانی را طی کند، زیرا صدایی که در یک مکان تولید می‌شود، موج صوتی را به وجود می‌آورد که در رسانای آکوستیکی به حرکت درمی‌آید. موج صوتی، یک نوسان فشار در حال حرکت است که در درون هر رسانای قابل ارتجاعی که ملکول‌های آن بتوانند به یکدیگر فشرده و از هم دور شوند، انتشار می‌یابد؛ مانند موجی که با پرتاب سنگ به درون یک دریاچه تولید می‌شود. ضربه سنگ به فاصله نسبتاً بزرگی انتقال پیدا می‌کند. ملکول‌های آب حرکت نمی‌کنند، بلکه این نوسان فشار است که به حرکت درمی‌آید.

صف افرادی که منتظر ورود به سینما هستند، مثال خوبی برای توصیف یک موج صوتی است. هنگامی که فرد جلوی صف حرکت می‌کند، یک «فضای خالی» بین او و نفر بعدی در صف ایجاد می‌شود (فاصله بین آنها افزایش پیدا می‌کند). از این رو

نفر دوم به جلو گام برمی‌دارد. در این صورت، یک فضای خالی بین نفر دوم و نفر سوم در صف ایجاد می‌شود و بنابراین، نفر سوم نیز به جلو حرکت می‌کند. سرانجام، آخرین نفر صف شروع به حرکت می‌کند. آخرین نفر تحت تأثیر حرکتی قرار گرفته که در ابتدای صف رخ داده است، زیرا نوسان فشار (فضای خالی درون صف) به انتهای صف انتقال یافته است، هرچند که هر فرد فقط کمی در صف جابه‌جا شده است. اما این مقایسه از جهتی اشکال دارد، زیرا در بیشتر صف‌ها افراد فقط به سمت جلو حرکت می‌کنند. برای آنکه مقایسه مناسبی با انتشار صوت صورت بگیرد، باید چنین تصور کنیم که نفر اول به عقب به سمت نفر دوم برمی‌گردد و این تراکم یا افزایش فشار (مثل فضای خالی) به سمت انتهای صف انتقال پیدا می‌کند.

شکل ۲-۱ یک موج فشار را در مکانی که در شکل ۱-۱ با علامت ستاره مشخص شده، نشان می‌دهد. محور افقی، گذر زمان و محور عمودی، میزان تجمع را (که در موج صوتی برابر است با فشار هوا) نشان می‌دهد. در زمان ۳، اُفتی ناگهانی در میزان تجمع صف دیده می‌شود، زیرا فرد شماره ۲ به جلو حرکت می‌کند و یک فاصله خالی در صف به جای می‌گذارد. در زمان ۴، تجمع صف دوباره عادی است، زیرا فرد شماره ۳ به جلو حرکت کرده و فاصله خالی به‌جامانده از فرد ۲ را پر می‌کند. در زمان ۱۰، یک افزایش ناگهانی در تجمع صف روی می‌دهد، زیرا فرد شماره ۲ به سمت عقب حرکت کرده و با فرد ۳ برخورد می‌کند. نمودار شکل ۲-۱ روشی برای بازنمایی امواج رقت و تراکم در حال حرکتی است که در شکل ۱-۱ نشان داده شد. شکل ۱-۱ را با فرض یک رسانای آکوستیکی واحد می‌توان از شکل ۲-۱ به دست آورد (البته به بحث پاراگراف بعدی در زمینه اتلاف انرژی صوتی نیز توجه کنید). نمودارهایی مانند نمودار شکل ۲-۱ معمولاً در آواشناسی آکوستیک به کار می‌روند، زیرا بازنمایی موج صوتی را، وقتی که به وسیله یک میکروفون تولید می‌شود، نشان می‌دهند؛ به عبارت دیگر، نوسانات دامنه را هنگامی که از یک نقطه معین در فضا می‌گذرند، نشان می‌دهند.

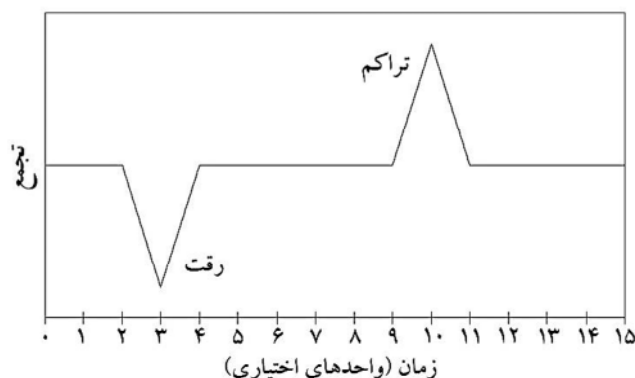
مثالی برای انتشار صوت

شکل ۱-۱ هفت نفر را (با عدد) نشان می‌دهد که برای دیدن یک نمایش در صف ایستاده‌اند. در زمان ۲، نفر اول به جلو حرکت کرده و یک فاصله در صف ایجاد می‌کند. سپس نفر دوم در زمان ۳، به جلو قدم برمی‌دارد و یک فاصله بین او و نفر سوم در صف باقی می‌ماند. فاصله خالی به سمت عقب صف حرکت می‌کند تا زمان ۸ که هر فرد یک قدم به جلو حرکت کرده است. در زمان ۹، نفر اول یک قدم به عقب برمی‌گردد و با نفر دوم برخورد می‌کند (با علامت x مشخص شده است). طبیعی است که نفر دوم در زمان ۱۰، از مسیر نفر اول خارج شده و با نفر سوم برخورد می‌کند. همان‌طور که در ابتدا فاصله خالی به سمت عقب صف حرکت می‌کند، حالا برخورد افراد نیز به سمت عقب صف به حرکت درمی‌آید تا در زمان ۱۵ که هر فرد به نقطه شروع حرکت خود برمی‌گردد.

می‌توان واژه‌های این مثال را با اصطلاحات انتشار صوت جایگزین کرد. بدین صورت که افراد ایستاده در صف متناظر با ملکول‌های هوا و گروه افراد، متناظر با یک رسانای آکوستیکی هستند. فاصله خالی بین افراد متوالی در صف، برابر با فشار منفی هوا یا رقت است و برخورد افراد، برابر با فشار مثبت هوا یا تراکم است. فشار صفر هوا (که در انتشار صوت، فشار اتمسفری است) همان فاصله طبیعی و اختیاری بین افراد ایستاده در یک صف است. جابه‌جایی اولیه نفر اول صف، مانند جابه‌جایی ملکول‌های هوای مجاور یک شاخک دیپازن (به عنوان مثال) است هنگامی که شاخک از ملکول‌ها دور می‌شود. جابه‌جایی نفر اول در زمان ۹، مانند حرکت مخالف یک شاخک دیپازن است وقتی که شاخک به ملکول‌ها نزدیک می‌شود.

زمان	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
		۱	۱	۱	۱	۱	۱	۱							
	۱		۲	۲	۲	۲	۲	۲	X	۱	۱	۱	۱	۱	۱
*	۲	۲		۳	۳	۳	۳	۳	۳	X	۲	۲	۲	۲	۲
	۳	۳	۳		۴	۴	۴	۴	۴	۴	X	۳	۳	۳	۳
	۴	۴	۴	۴		۵	۵	۵	۵	۵	۵	X	۴	۴	۴
	۵	۵	۵	۵	۵		۶	۶	۶	۶	۶	۶	X	۵	۵
	۶	۶	۶	۶	۶	۶		۷	۷	۷	۷	۷	۷	X	۶
	۷	۷	۷	۷	۷	۷	۷								۷

شکل ۱-۱. حرکت موجی در درون صفی هفت نفره برای ورود به یک سالن نمایش. زمان در بالای نمودار از ابتدا (زمان ۱) تا انتها (زمان ۱۵) به صورت واحدهای اختیاری نشان داده شده است.



شکل ۱-۲. موج فشار حرکت موجی که در شکل ۱-۱ نشان داده شد. محور افقی، زمان و محور عمودی فاصله بین افراد را نشان می‌دهد.

امواج صوتی، همان‌طور که از هوا (یا هر رسانای آکوستیکی دیگری) می‌گذرند، انرژی از دست می‌دهند، زیرا حرکت ملکول‌های هوا مستلزم صرف انرژی است. شاید شما مشابه این پدیده را هنگامی که در یک صف طولانی ایستاده‌اید، ملاحظه کرده باشید. اگر نفر اول صف به سمت جلو قدم بردارد و سپس به سمت عقب برگردد، فقط عده کمی در جلوی صف تحت تأثیر این جلو و عقب‌رفتن او قرار می‌گیرند، زیرا افراد عقب‌تر صف اینرسی دارند و قبل از آنکه در واقع، در پاسخ به تغییر جابه‌جا شوند، اندکی تغییر فشار (فاصله بین افراد) را تحمل می‌کنند. بنابراین، آشفتگی در جلوی یک صف طولانی ممکن است هیچ تأثیری بر افراد انتهای آن نداشته باشد. از سویی ممکن است افراد داخل صف وول بخورند. در این حالت، اگر جابه‌جایی انتشار یافته به سمت عقب صف کوچک باشد، ممکن است تشخیص تفاوت بین آن با جابه‌جایی ناشی از وول‌خوردن افراد (نسبت سیگنال به نوفه^۱) دشوار باشد. سرعت اتلاف انرژی صوتی در هوا با اتلاف جابه‌جایی در یک صف متفاوت است، زیرا صدا از یک منبع صوتی در سه بعد (در یک محیط) منتشر می‌شود. یعنی تعداد ملکول‌هایی که موج صوتی سبب جابه‌جاشدن آنها می‌شود، همان‌طور که موج از منبع صوتی منتشر می‌گردد، به‌میزان قابل توجهی افزایش پیدا

1. signal to noise ratio

می‌کند. بنابراین، همان‌طور که موج از منبع صوتی به بیرون منتشر می‌شود، مقدار انرژی موجود برای حرکت ملکول‌ها در واحد سطح محیط کاهش می‌یابد. در نتیجه، میزان جابه‌جایی ملکول‌ها نیز به‌عنوان تابعی از فاصله، با منبع صوتی (با توان ۳) کاهش پیدا می‌کند. به‌همین دلیل است که خوانندگان گروه‌های هوی‌متال^۱، میکروفون را مستقیماً جلوی لب‌هایشان می‌گیرند، زیرا در غیر این صورت، صدای آنها در سروصدای محیط گم می‌شود. به‌همین دلیل شما نیز باید هنگام ضبط یک نمونه گفتاری، میکروفون را نزدیک دهان گویشور قرار دهید (گرچه بسیار مهم است که برای جلوگیری از تأثیر نوفه‌های انفجار آوایی مثل [p] و غیره، میکروفون را مجاور لب‌های گویشور، و نه مستقیماً روبروی آن، قرار دهید).

۳-۱. اقسام صوت

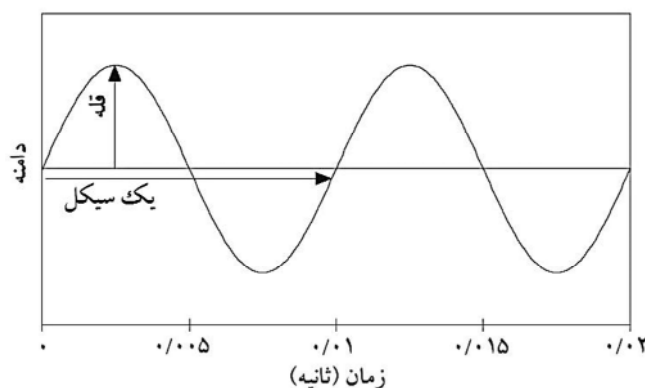
امواج صوتی به دو دسته تقسیم می‌شوند: متناوب و نامتناوب. امواج متناوب الگویی دارند که در فواصل زمانی منظم تکرار می‌شود. این امواج به دو نوع ساده و مرکب تقسیم می‌شوند.

۱-۳-۱. امواج متناوب ساده

امواج متناوب ساده را امواج سینوسی نیز می‌نامند. این امواج از یک حرکت هماهنگ^۲ ساده مثل نوسان یک پاندول به وجود می‌آیند. ما انسان‌ها تنها هنگامی امواجی تقریباً شبیه به امواج متناوب ساده را در گفتارمان تولید می‌کنیم که خیلی کم‌سن‌وسال هستیم. ارتعاش تارهای صوتی کودکان تقریباً سینوسی است و معمولاً ارتعاش تارهای صوتی زنان سینوسی‌تر از مردان است. علی‌رغم این واقعیت که امواج متناوب ساده به‌ندرت در گفتار تولید می‌شوند، اما بسیار مهم هستند، زیرا امواج مرکب را می‌توان به‌صورت ترکیباتی از امواج سینوسی توصیف کرد. برای تعریف یک موج سینوسی، دانستن سه ویژگی آن لازم است. این ویژگی‌ها در شکل‌های ۱-۳ و ۱-۴ نشان داده شده‌اند.

1. heavy metal bands

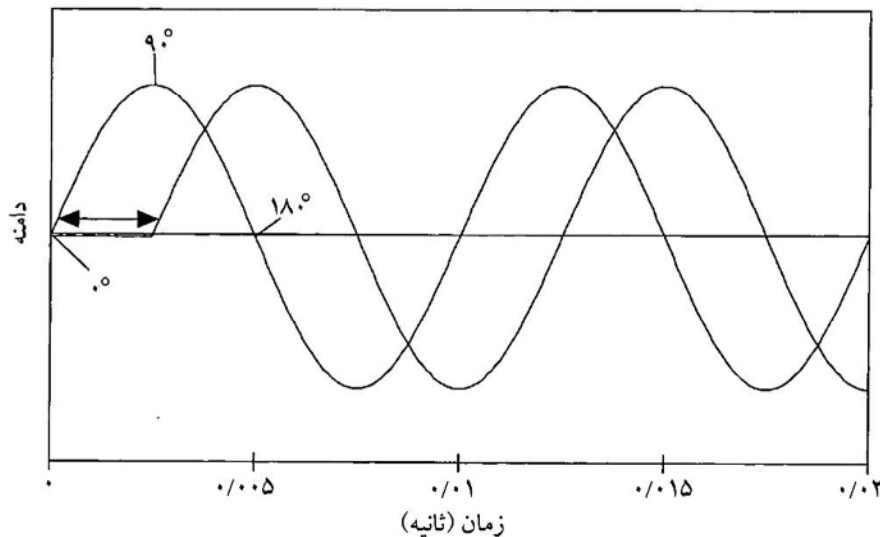
2. harmonic



شکل ۱-۳. یک موج سینوسی ۱۰۰ هرتزی که دیرش یک سیکل (دوره تناوب) و قله دامنه آن مشخص شده است.

اولین ویژگی یک موج متناوب ساده، **فرکانس** آن است: تعداد دفعاتی که الگوی سینوسی در واحد زمان (بر روی محور افقی) تکرار می‌شود. هر تکرار الگوی سینوسی یک **سیکل** نامیده می‌شود و مدت زمان یک سیکل، برابر با **دوره تناوب** آن است. فرکانس را می‌توان به صورت تعداد سیکل‌ها در واحد زمان تعریف کرد که در این صورت به طور قراردادی واحد آن **هرتز** (با علامت اختصاری Hz) است. بنابراین، برای به دست آوردن فرکانس یک موج سینوسی برحسب Hz (تعداد سیکل‌ها در ثانیه)، یک ثانیه را بر دوره تناوب (مدت زمان یک سیکل) تقسیم می‌کنند. به عبارت دیگر، فرکانس برحسب هرتز مساوی است با $1/T$ ، که T دوره تناوب است برحسب ثانیه. برای مثال، موج سینوسی شکل ۱-۳، یک سیکل را در 0.01 ثانیه کامل می‌کند. تعداد سیکل‌هایی که این موج در یک ثانیه می‌تواند کامل کند، ۱۰۰ سیکل است (یعنی یک ثانیه تقسیم بر مدت زمانی که یک سیکل در ثانیه طول می‌کشد یا $\frac{1}{0.01} = 100$).

دومین ویژگی یک موج متناوب ساده، **دامنه** آن است: قله انحراف یک نوسان فشار از فشار نرمال اتمسفری. دامنه یک موج فشار صوتی بر روی محور عمودی نشان داده می‌شود.

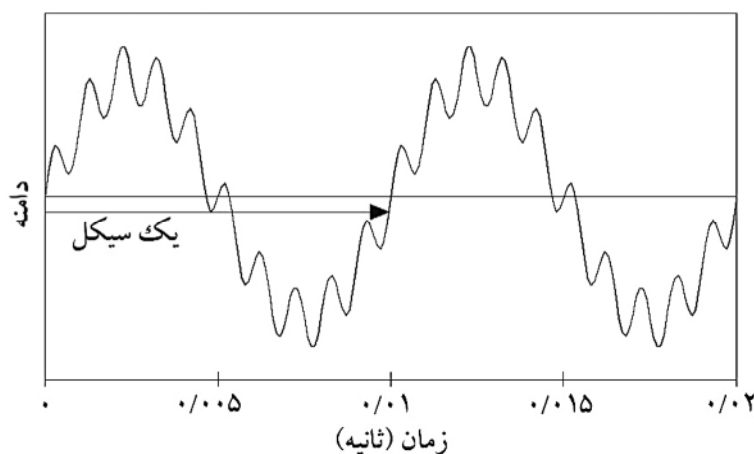


شکل ۴-۱. دو موج سینوسی با فرکانس و دامنه یکسان اما با اختلاف فاز ۹۰ درجه.

سومین ویژگی امواج سینوسی، فاز آنهاست: زمان‌بندی موج صوتی نسبت به یک نقطه مرجع. شما می‌توانید با به‌دست آوردن مقادیر دامنه از مجموعه‌ای از مثلث‌های قائمه که در درون یک دایره قرار دارند، یک موج سینوسی رسم کنید (برای رسم موج سینوسی به تمرین ۴ در پایان این فصل رجوع کنید). یک بار حرکت دور دایره برابر است با یک موج سینوسی بر روی کاغذ. بنابراین، می‌توانیم نقاط یک موج سینوسی را بر اساس درجه چرخش حول یک دایره تعیین کنیم. این نقاط در شکل ۴-۱ نشان داده شده است. در این شکل، نقطه شروع هر دو موج سینوسی در صفر درجه سیکل سینوسی است. در هر دو، قله دامنه در ۹۰ درجه و در هر دو، نقطه عبور از صفر دامنه در جهت رو به پایین (جهت منفی) در ۱۸۰ درجه است. قله منفی این دو موج در ۲۷۰ درجه و نقطه پایانی سیکل در ۳۶۰ درجه است. اما این دو موج سینوسی که دامنه و فرکانس دقیقاً یکسانی دارند، ممکن است به لحاظ زمان‌بندی نسبی یا فاز با یکدیگر اختلاف داشته باشند. در این مثال، دو موج با یکدیگر ۹۰ درجه اختلاف فاز دارند.

۱-۳-۲. امواج متناوب مرکب

امواج متناوب مرکب شبیه به امواج متناوب ساده هستند، زیرا الگوی شکل موج آنها تکرار می‌شود و بنابراین، سیکل دارند. اما امواج متناوب مرکب از حداقل دو موج سینوسی ساده ترکیب یافته‌اند. برای مثال، به موج شکل ۱-۵ توجه کنید. این موج نیز مانند امواج سینوسی ساده شکل‌های ۱-۳ و ۱-۴، یک سیکل را در 0.01 ثانیه (یا 10 میلی‌ثانیه) کامل می‌کند. اما این موج مؤلفه^۱ دیگری نیز دارد که ده سیکل را در همین مدت زمان کامل می‌کند. به موج‌های کوچک روی شکل موج نگاه کنید. می‌توانید ده قله کوچک مثبت بر روی یک سیکل آن بشمارید. هر یک از این قله‌های کوچک، یک سیکل مؤلفه دیگر موج مرکب را نشان می‌دهد. من این موج را از ترکیب یک موج سینوسی ساده 100 هرتزی با یک موج سینوسی ساده 1000 هرتزی (با دامنه کمتر) تولید کردم. بنابراین، از ترکیب یک موج 1000 هرتزی با یک موج 100 هرتزی، یک موج متناوب مرکب تولید می‌شود. نرخی که در آن الگوی مرکب تکرار می‌شود، فرکانس پایه (با علامت اختصاری F_0) نامیده می‌شود.



شکل ۱-۵. یک موج متناوب مرکب که از یک موج سینوسی 100 هرتزی و یک موج سینوسی 1000 هرتزی ترکیب یافته است. در این موج، یک سیکل فرکانس پایه (F_0) مشخص شده است.

فرکانس پایه و بزرگ‌ترین مقسوم‌علیه مشترک

فرکانس پایه موجی که در شکل ۱-۵ نشان داده شد، ۱۰۰ هرتز است. این موج یک مؤلفه سینوسی ۱۰۰ هرتزی نیز دارد. فرکانس پایه یک موج مرکب عبارت است از بزرگ‌ترین مقسوم‌علیه مشترک فرکانس امواج سینوسی سازنده آن. برای مثال، فرکانس پایه (F_0) یک موج مرکب حاصل از مؤلفه‌های ۴۰۰ هرتزی و ۵۰۰ هرتزی، ۱۰۰ هرتز است. شما می‌توانید خودتان آن را با رسم موج متناوب مرکب حاصل از افزودن امواج سینوسی ۴۰۰ هرتزی و ۵۰۰ هرتزی ببینید. می‌توانیم از موج سینوسی شکل ۱-۳ برای رسم این نمودار استفاده کنیم. مراحل رسم این موج به صورت زیر است:

۱. یک کاغذ رسم نمودار تهیه کنید.
۲. دوره تناوب موج سینوسی ۴۰۰ هرتزی را حساب کنید. از آنجا که فرکانس برابر است با یک تقسیم بر دوره تناوب ($f = 1/T$)، پس دوره تناوب برابر است با یک تقسیم بر فرکانس ($T = 1/f$). بنابراین، دوره تناوب یک موج سینوسی ۴۰۰ هرتزی، ۰/۰۰۲۵ ثانیه است که برحسب میلی ثانیه (یک هزارم ثانیه) برابر با ۲/۵ میلی ثانیه ($0/0025 \times 1000$) است.
۳. دوره تناوب موج سینوسی ۵۰۰ هرتزی را نیز محاسبه کنید.
۴. اکنون می‌توانیم جدولی از مقادیر لازم برای رسم امواج سینوسی ۴۰۰ هرتزی و ۵۰۰ هرتزی را به دست آوریم. برای این کار، ابتدا بر روی محور زمان در شکل ۱-۳ تعدادی نقطه جدید برای هر یک از امواج ۴۰۰ هرتزی و ۵۰۰ هرتزی علامت‌گذاری کنید. برای موج سینوسی ۴۰۰ هرتزی، بر روی محور زمان به جای نقطه ۰/۰۱ ثانیه، زمان ۲/۵ میلی ثانیه را علامت‌گذاری کنید، زیرا یک سیکل این موج، ۲/۵ میلی ثانیه طول می‌کشد. به جای نقطه ۰/۰۰۵ ثانیه نیز زمان ۱/۲۵ میلی ثانیه را علامت‌گذاری کنید. قله موج سینوسی ۴۰۰ هرتزی در زمان ۰/۶۲۵ میلی ثانیه و دره آن در ۱/۸۷۵ میلی ثانیه است. بدین ترتیب، جدول مقادیر زمان و دامنه موج ۴۰۰ هرتزی به دست می‌آید. (فرض می‌کنیم که دامنه قله ۱ و دامنه دره ۱-)

است و مقدار دامنه در زمان $3/125$ میلی‌ثانیه برابر با مقدار دامنه قله در سیکل بعدی است):

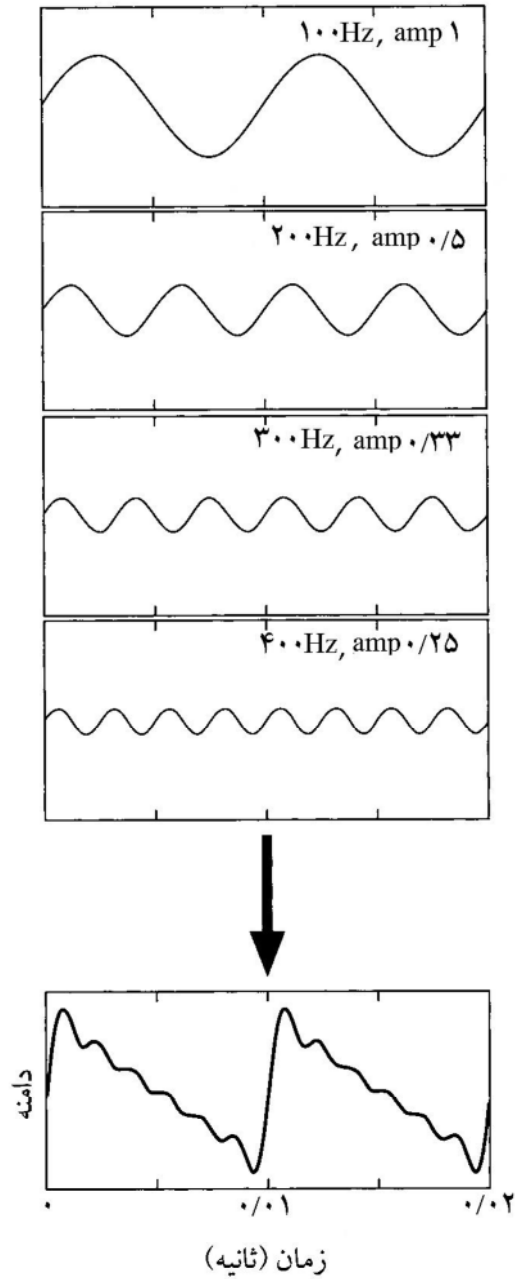
میلی‌ثانیه	۰	$0/625$	$1/25$	$1/875$	$2/5$	$3/125$
دامنه	۰	۱	۰	-۱	۰	۱

فاصله بین نقاط متوالی موج (با فاصله ۹۰ درجه از یکدیگر) $0/625$ میلی‌ثانیه است. در موج ۵۰۰ هرتزی فاصله بین نقاط، $0/5$ میلی‌ثانیه است. ۵. حالا روی کاغذ رسم نمودار، ۲۰ میلی‌ثانیه را با فاصله ۱ میلی‌ثانیه از یکدیگر علامت‌گذاری کنید. همچنین، مقیاس دامنه‌ای از ۱ تا -۱ با فاصله‌ای در حدود یک اینچ علامت‌گذاری کنید.

۶. امواج سینوسی ۴۰۰ هرتزی و ۵۰۰ هرتزی را با علامت‌گذاری نقاط تلاقی مقادیر زمان و دامنه بر روی نمودار رسم کنید. برای مثال، اولین نقطه موج ۴۰۰ هرتزی، در زمان صفر میلی‌ثانیه و دامنه صفر خواهد بود. دومین نقطه در زمان $0/625$ میلی‌ثانیه و دامنه ۱ است و به همین ترتیب، الی آخر. شما ممکن است بخواهید جدول را تا ۲۰ میلی‌ثانیه ادامه دهید (من برای حفظ زمان دقیق موج ۴۰۰ هرتزی، فقط تا $3/125$ میلی‌ثانیه ادامه دادم). وقتی همه نقاط موج ۴۰۰ هرتزی را علامت‌گذاری کردید، آنها را به صورت یک موج سینوسی به هم وصل کنید. سپس، موج سینوسی ۵۰۰ هرتزی را به همین شیوه و با استفاده از همین محورهای زمان و دامنه رسم کنید. در نهایت، شما باید نموداری با امواج سینوسی روی هم افتاده، چیزی شبیه به شکل ۱-۶ به دست آورید.

۷. حالا مقادیر دو موج را به یکدیگر بیفزایید. در هر $0/5$ میلی‌ثانیه فاصله، مقادیر دامنه دو موج سینوسی را برای به دست آوردن مقدار دامنه موج متناوب مرکب جدید با هم جمع کنید و سپس شکل موج نهایی را رسم کنید.

اکنون به موج متناوب مرکب حاصل از ترکیب موج سینوسی ۴۰۰ هرتزی با موج سینوسی ۵۰۰ هرتزی نگاه کنید. آیا فرکانس پایه آن ۱۰۰ هرتز است؟ اگر چنین است شما باید دو سیکل کامل در طول ۲۰ میلی‌ثانیه موج مرکب ببینید و الگوی شکل موج از ۱۰ میلی‌ثانیه تا ۲۰ میلی‌ثانیه باید کاملاً همانند الگویی باشد که از صفر میلی‌ثانیه تا ۱۰ میلی‌ثانیه می‌بینید.



شکل ۱-۶. شکل موج یک موج متناوب مرکب که تقریباً شبیه به یک موج دندان‌اره‌ای است و چهار موج سینوسی با فرکانس پایین که برای تولید آن با یکدیگر ترکیب شده‌اند.

شکل ۱-۶ موج مرکب دیگری را نشان می‌دهد (و چهار موج سینوسی ساده که برای تولید آن با هم ترکیب شده‌اند). شکل موج آن تقریباً شبیه به الگوی دندان‌اره‌ای^۱ است. برخلاف مثال قبلی، نمی‌توان فقط با نگاه کردن به شکل موج مرکب، امواج سینوسی سازنده آن را تعیین کرد. توجه کنید که چگونه هر چهار مؤلفه سینوسی آن، دارای قله‌های مثبت در ابتدای سیکل موج مرکب و قله‌های منفی در انتهای سیکل است. قله‌ها به یکدیگر افزوده می‌شوند و یک قله تند در ابتدای سیکل و یک دره تند در انتهای آن به وجود می‌آورند و در بقیه سیکل، تمایل به خنثی کردن یکدیگر دارند. در موج مرکب قله‌های مجزایی که نشان‌دهنده سیکل‌های مؤلفه‌های موج باشند، قابل مشاهده نیست. با این همه، موج مرکب حاصل افزودن امواج ساده به یکدیگر است.

اکنون ببینیم چگونه می‌توان مؤلفه‌های فرکانسی یک موج متناوب مرکب را بازنمایی کرد. در واقع هدف ما یافتن شیوه‌ای است برای نشان دادن مؤلفه‌های سینوسی یک موج مرکب وقتی که به آسانی در شکل موج قابل مشاهده نیستند. یک روش این است که فرکانس و دامنه امواج سینوسی سازنده آن را به صورت زیر فهرست کنیم:

فرکانس (هرتز)	۱۰۰	۲۰۰	۳۰۰	۴۰۰	۵۰۰
دامنه	۱	۰/۵	۰/۳۳	۰/۲۵	۰/۲

در اینجا به مبحث پیچیده دامنه نمی‌پردازم. دامنه امواج سینوسی را می‌توان با مقیاس‌های اندازه‌گیری مختلفی برحسب بزرگی^۲، شدت^۳ یا بلندی درکی^۴ آن توصیف کرد (برای مطالب بیشتر به فصل ۴ رجوع کنید). در این فصل، بزرگی امواج صوتی را به صورت نسبی نشان می‌دهم و به همین دلیل نیازی به ارائه واحدهای اندازه‌گیری دامنه نیست؛ به این صورت که مؤلفه ۲۰۰ هرتزی دامنه‌ای دارد که برابر با نصف بزرگی مؤلفه ۱۰۰ هرتزی است و به همین ترتیب، الی آخر.

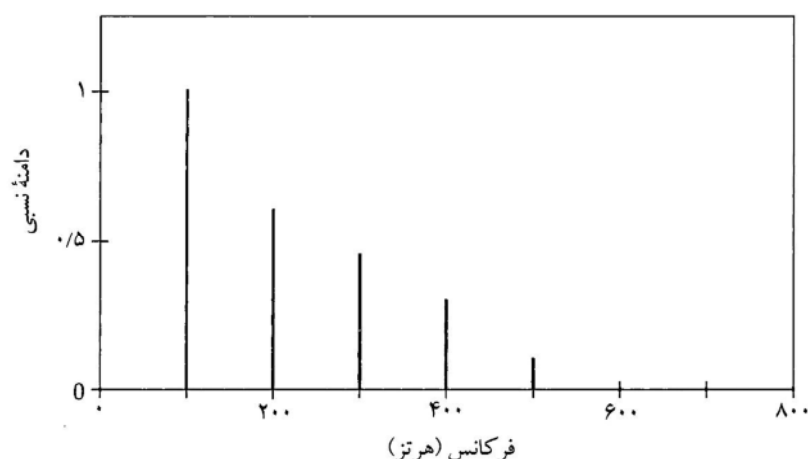
1. sawtooth pattern

2. magnitude

3. intensity

4. perceived loudness

شکل ۱-۷ نموداری را بر اساس مقادیر بالا نشان می‌دهد. در این نمودار محور افقی، بازنمون فرکانس و محور عمودی، بازنمون دامنه است. نمایش نموداری مؤلفه‌های فرکانسی، بهترین شیوه برای نشان دادن مؤلفه‌های سینوسی یک موج متناوب مرکب است، زیرا امواج مرکب اغلب از مؤلفه‌های فرکانسی بسیار زیادی ترکیب یافته‌اند که نشان دادن آنها در جدول عملی نیست. نمودار مقادیر دامنه مؤلفه‌های سینوسی یک موج مرکب در مقابل فرکانس آنها، طیف قدرت^۱ آن موج مرکب نامیده می‌شود.



شکل ۱-۷. یک بازنمایی نموداری از مقادیر فرکانس و دامنه امواج سینوسی موج مرکبی که در شکل ۱-۶ نشان داده شد.

اهمیت این موضوع که امواج متناوب مرکب را می‌توان با افزودن امواج سینوسی ساده به یکدیگر ساخت، در چیست؟ در واقع، می‌توان شمار نامحدود و متنوعی از امواج مرکب را با ترکیب امواج سینوسی ساده‌ای که فرکانس، دامنه و فازهای مختلفی دارند، تولید نمود. یکی از ویژگی‌های امواج صوتی این است که می‌توان هر موج مرکب را به امواج سینوسی سازنده آن تجزیه کرد. به عبارت دیگر، هر موج مرکب را می‌توان به مجموعه‌ای از امواج سینوسی ساده که فرکانس، دامنه و فاز

1. power spectrum

معینی دارند، تجزیه نمود. این ویژگی امواج صوتی را بر اساس نام ریاضیدان قرن هفدهم، «ژوزف فوریه» که آن را کشف کرد، قضیه فوریه می‌نامند.

در تجزیه فوریه^۱ یک موج متناوب مرکب را که از تعدادی مؤلفه اختیاری تشکیل شده، انتخاب کرده و مقادیر فرکانس، دامنه و فاز هر یک از مؤلفه‌ها را استخراج می‌کنیم. حاصل تجزیه فوریه یک طیف قدرت است شبیه به طیفی که در شکل ۱-۷ نشان داده شد. (از آنجا که فاز مؤلفه‌های موج، تأثیر جزئی بر درک صدا دارد، از آن صرف نظر می‌کنیم).

۱-۳-۳. امواج نامتناوب

امواج نامتناوب برخلاف امواج متناوب ساده یا مرکب، الگوی تکراری متناوبی ندارند. این امواج یا یک شکل موج تصادفی دارند یا الگویی که تکرار نمی‌شود. موجی که مشخصه آن نوسان فشار تصادفی است، نوفه سفید نامیده می‌شود و صدای آن شبیه به پارازیت رادیویی یا وزش باد در لابه‌لای برگ درختان است. هرچند که نوفه سفید یک موج متناوب نیست، اما می‌توان با تجزیه فوریه آن را تجزیه کرد. برخلاف تجزیه فوریه امواج متناوب، طیف نوفه سفید قله‌های تند ندارد، بلکه همه مؤلفه‌های فرکانسی محتمل آن، دامنه یکسان دارند (طیف آن مسطح^۲ است). نوفه سفید نیز مانند موج سینوسی یک مفهوم انتزاعی است، هرچند که بسیاری از صداهای طبیعی مثل صدای باد یا آواهای سایشی [s] یا [f] شبیه به نوفه سفید هستند.

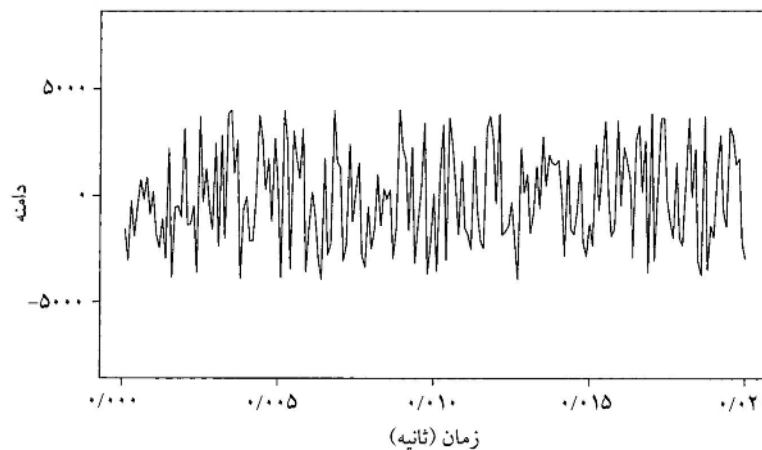
شکل‌های ۱-۸ و ۱-۹ به ترتیب، شکل موج آکوستیکی و طیف قدرت یک نمونه نوفه سفید را نشان می‌دهند. ملاحظه کنید که شکل موج آکوستیکی شکل ۱-۸ نامتناوب است و هیچ الگوی تکراری قابل تشخیصی در آن دیده نمی‌شود. همچنین، به طیف شکل ۱-۹ نگاه کنید که در بالا مسطح است. همان‌طور که در فصل ۳ (در پردازش دیجیتال سیگنال) نیز خواهیم دید، تجزیه فوریه قطعه کوتاهی از شکل موج (که «پنجره تجزیه»^۳ نامیده می‌شود)، سبب بروز اشتباهاتی در طیف

1. Fourier analysis

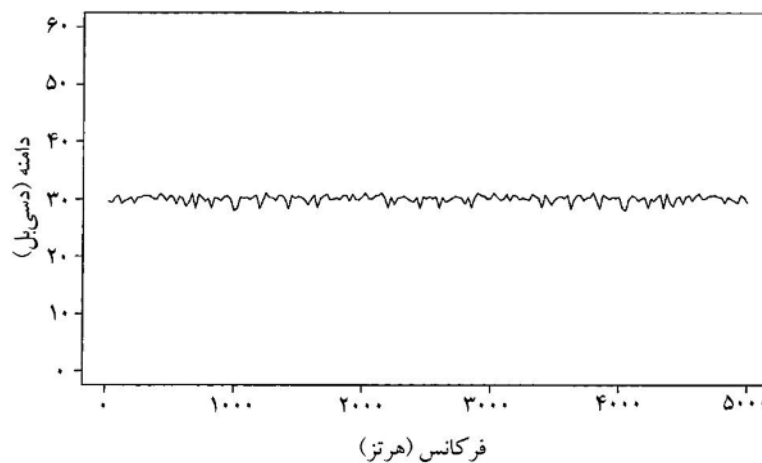
2. flat

3. analysis window

به دست آمده می شود. به همین دلیل، این طیف، قله ها و دره هایی دارد، هر چند که طبق نظریه، طیف نوفه سفید باید مسطح باشد.



شکل ۸-۱ یک قطعه ۲۰ میلی ثانیه ای از شکل موج آکوستیکی یک نوفه سفید. دامنه در هر نقطه از زمان تصادفی است.

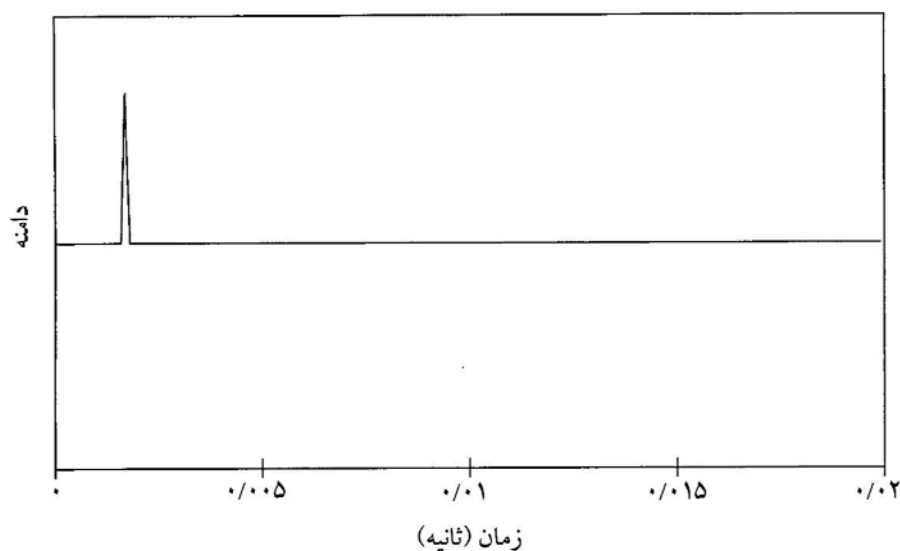


شکل ۹-۱. طیف قدرت نوفه سفیدی که در شکل ۸-۱ نشان داده شد.

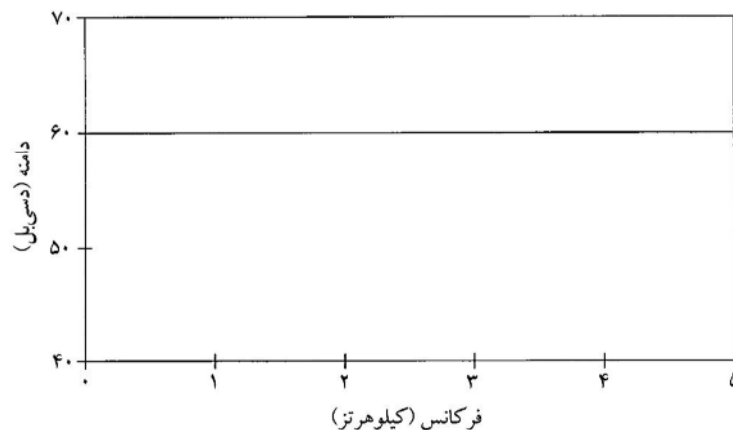
نوع اصلی دیگر امواج نامتناوب، امواج گذرا^۱ نامیده می شوند. این امواج که اقسام مختلفی مثل صداهای انفجاری و به هم خوردن اشیاء را شامل می شوند، یک

1. transient

نوسان فشار ناگهانی تولید می‌کنند که در بُعد زمان تداوم یا تکرار نمی‌یابند. صدای به‌هم‌خوردن درها، ترکیدن بادکنک و کلیک‌های الکتریکی همه، امواج گذرا هستند. می‌توان با استفاده از تجزیه فوریه، امواج گذرا را نیز مانند نوفه سفید به مؤلفه‌های طیفی تجزیه کرد. شکل ۱-۱۰ نمونه ایده‌آل یک سیگنال گذرا را نشان می‌دهد. در سیگنال موج گذرا، انرژی فقط در یک نقطه از زمان مشاهده می‌شود و در سایر زمان‌ها فشار مساوی صفر است. این نمونه ایده‌آل، پالس نامیده می‌شود. شکل امواج گذرای طبیعی تقریباً شبیه به یک پالس است، اما معمولاً با نوسانی نسبتاً پیچیده‌تر. شکل ۱-۱۱ طیف قدرت شکل موج پالس شکل ۱-۱۰ را نشان می‌دهد. طیف آن مثل طیف نوفه سفید مسطح است. مسطح بودن طیف در شکل ۱-۱۱ واضح‌تر از شکل ۱-۹ است، زیرا «پالسی بودن» شکل موج پالس تنها به یک نقطه در زمان بستگی دارد، در حالی که «نوفه سفید بودن» شکل موج نوفه سفید به هر نقطه از زمان ارتباط پیدا می‌کند. بنابراین، چون تجزیه فوریه تقریباً برای نمونه کوچکی از یک شکل موج کاربرد دارد، طیف نوفه سفید به دقت طیف پالس به دست نمی‌آید.



شکل ۱-۱۰. شکل موج آکوستیکی یک موج گذرا (یک پالس).



شکل ۱-۱۱. طیف قدرت سیگنال گذرای که در شکل ۱-۱۰ نشان داده شد.

۴-۱. صافی‌های^۱ آکوستیکی

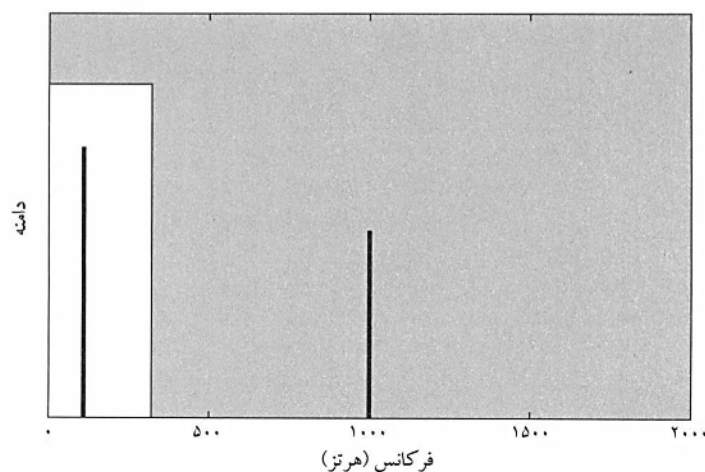
همه ما با طرز کار صافی‌ها آشنا هستیم. به عنوان مثال، شما با استفاده از یک صافی کاغذی، دانه‌های قهوه را از قهوه‌تان یا با یک چای صاف‌کن برگ‌های چای را از چایی‌تان جدا می‌کنید. این مثال‌های روزمره برخی از ویژگی‌های مهم صافی‌های آکوستیکی را نشان می‌دهند. مثلاً تفاوت اصلی بین یک قهوه‌صاف‌کن با یک چای‌صاف‌کن در این است که چای‌صاف‌کن ذرات بزرگ‌تری را عبور می‌دهد، اما قهوه‌صاف‌کن ذرات ریزتری را جذب می‌کند. پس تفاوت بین این دو صافی در اندازه ذراتی است که عبور می‌دهند.

صافی آکوستیکی به جای عبور یا جذب ذرات با اندازه‌های مختلف مثل یک قهوه‌صاف‌کن، مؤلفه‌های صوتی با فرکانس‌های مختلف را عبور می‌دهد یا جذب می‌کند. به عنوان مثال، یک صافی آکوستیکی پایین‌گذر^۲، مؤلفه‌های دارای فرکانس بالای یک موج را جذب می‌کند و مؤلفه‌های دارای فرکانس پایین را عبور می‌دهد. پیش‌تر، تفاوت بین امواج متناوب ساده با امواج مرکب، با افزودن یک موج سینوسی ۱۰۰۰ هرتزی به یک موج سینوسی ۱۰۰ هرتزی برای تولید یک موج مرکب شرح داده شد. با یک صافی پایین‌گذر که برای مثال، همه مؤلفه‌های فرکانسی بالای ۳۰۰

۱. در متن، هر دو معادل «صافی» و «فیلتر» برای filter به کار رفته است.

هرتز را حذف می‌کند، می‌توانیم موج ۱۰۰۰ هرتزی را از موج مرکب حذف کنیم. همان‌طور که یک قهوه‌صاف‌کن ذرات ریز را عبور می‌دهد و ذرات درشت را جذب می‌کند، یک صافی آکوستیکی پایین‌گذر نیز به مؤلفه‌های دارای فرکانس پایین اجازه عبور می‌دهد، اما مؤلفه‌های دارای فرکانس بالا را متوقف می‌سازد.

شما می‌توانید عمل یک صافی پایین‌گذر را با نمایش طیف تابع پاسخ^۱ آن به تصویر بکشید. برای مثال، شکل ۱-۱۲ یک صافی پایین‌گذر را نشان می‌دهد که فرکانس قطع^۲ آن در ۳۰۰ هرتز است. قسمتی از طیف که سفیدرنگ است، نوارِ گذر نامیده می‌شود، زیرا انرژی صوتی در این محدوده فرکانسی از صافی عبور می‌کند، اما قسمتی از طیف که خاکستری‌رنگ است، نوارِ دفع نامیده می‌شود، زیرا انرژی صوتی در این ناحیه به وسیله صافی، دفع یا بلوکه می‌شود. بنابراین، در یک موج مرکب با مؤلفه‌های ۱۰۰ هرتزی و ۱۰۰۰ هرتزی، مؤلفه ۱۰۰ هرتزی عبور می‌کند، اما مؤلفه ۱۰۰۰ هرتزی متوقف می‌شود. به همین صورت، یک صافی بالاگذر^۳ مؤلفه‌های فرکانس پایین یک موج را جذب می‌کند و مؤلفه‌های فرکانس بالا را عبور می‌دهد. نمایش طیفی تابع پاسخ یک صافی بالاگذر نشان می‌دهد که مؤلفه‌های فرکانس پایین به وسیله صافی حذف می‌شوند و مؤلفه‌های فرکانس بالا عبور می‌کنند.



شکل ۱-۱۲. طیف یک صافی پایین‌گذر.

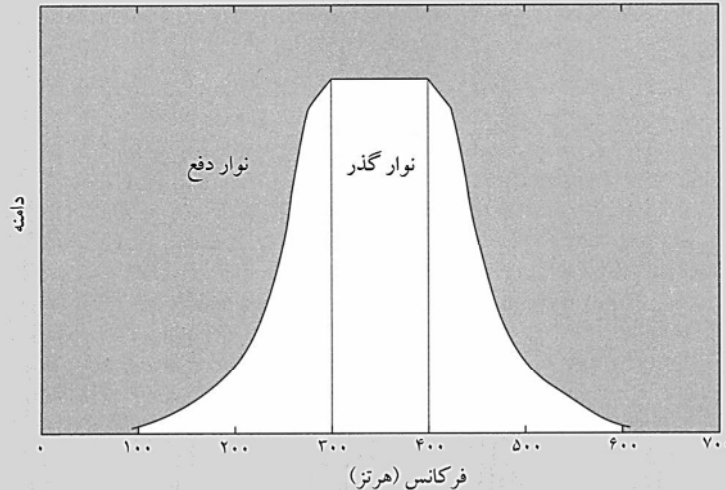
1. response function

2. cutoff frequency

3. high-pass filter

شیب صافی

در صافی پایین‌گذر شکل ۱-۱۲، مرز بسیار تندی در ۳۰۰ هرتز بین فرکانس‌هایی که از صافی عبور می‌کنند با آنهایی که دفع می‌شوند، وجود دارد. از آنجا که شیب خط عمودی که نوار گذر را از نوار دفع جدا می‌کند بی‌نهایت تند است، صافی عملکرد یکسانی بر هر مؤلفه فرکانسی زیر (یا بالای) فرکانس قطع دارد. در زندگی واقعی، صافی‌های آکوستیکی چنین مرزهای تندی ندارند. به‌طور مثال، مرز بین نوار عبور و نوار دفع معمولاً تا چند محدوده فرکانسی گسترش پیدا می‌کند (مثل صافی نوارگذر^۱ شکل ۱-۱۳) و بلافاصله ایجاد نمی‌شود (مثل صافی پایین‌گذر شکل ۱-۱۲). صافی که شیب بسیار تندی دارد، مثل چای صاف‌کنی است که روزنه‌های آن کاملاً هم‌اندازه است، اما صافی که شیب ملایمی دارد، مانند چای صاف‌کنی است که روزنه‌های آن اندازه‌های متفاوتی دارند؛ برخی ذرات به‌وسیله روزنه‌های کوچک‌تر جذب می‌شوند، اما همین ذرات اگر به روزنه‌های بزرگ‌تر برخورد کنند، از آنها عبور می‌کنند.



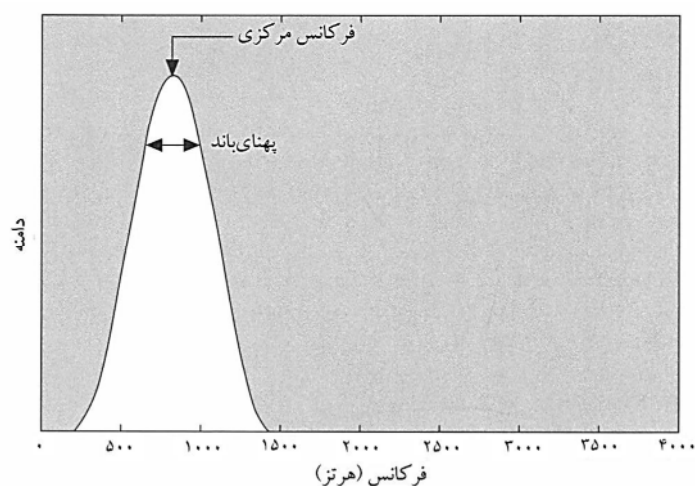
شکل ۱-۱۳. نمونه یک صافی نوار گذر. توجه کنید که صافی در هر دو طرف نوار گذر، حاشیه^۲ دارد.

صافی‌های نوارگذر مهم هستند، زیرا می‌توانیم برخی جنبه‌های تولیدی گفتار و شنوایی را بر اساس عمل این صافی‌ها مدل‌سازی کنیم. برخلاف صافی‌های پایین‌گذر یا بالاگذر که

1. band-pass filter 2. skirt

تنها یک فرکانس قطع دارند، صافی‌های نوارگذر (همان‌طور که شکل ۱-۱۳ نشان می‌دهد) دو فرکانس قطع دارند، یکی، قطع پایین نوار گذر و دیگری، قطع بالای نوار گذر. یک صافی نوارگذر مثل ترکیب یک صافی پایین‌گذر با یک صافی بالاگذر است، به نحوی که فرکانس قطع صافی پایین‌گذر بالاتر از فرکانس قطع صافی بالاگذر است.

هرگاه فرکانس قطع بالا و پایین یک صافی نوارگذر مساوی یکدیگر باشد، صافی حاصل را می‌توان بر اساس فرکانس مرکزی و پهنای باند^۱ (که به وسیله شیب‌های صافی تعیین می‌شود)، تعریف نمود. پهنای باند عبارت است از پهنای قله صافی (برحسب هرتز) به گونه‌ای که نیمی از انرژی آکوستیکی صافی در درون این ناحیه است. یعنی اگر کل ناحیه زیر منحنی صافی را در نظر بگیریم، پهنای باند محدوده فرکانسی حول فرکانس مرکزی است که دربرگیرنده نیمی از کل ناحیه زیر منحنی است. در عمل، پهنای باند که نیمی از قدرت صافی را دارد، با اندازه‌گیری دامنه در فرکانس مرکزی صافی و سپس یافتن پهنایی از صافی که دامنه آن سه دسی‌بل کمتر از دامنه قله است، به دست می‌آید (برای تعریف دسی‌بل به فصل ۴ رجوع کنید). این نوع خاص صافی نوارگذر که در شکل ۱-۱۴ نشان داده شده، در آواشناسی آکوستیک مهم است، زیرا برای مدل‌سازی صافی دستگاه گفتار و سیستم شنوایی به کار می‌رود.



شکل ۱-۱۴. توصیف یک صافی نوارگذر بر اساس فرکانس مرکزی و پهنای باند آن.

منابعی برای مطالعه بیشتر

Fry, D. B. (1979) *The Physics of Speech*, Cambridge: Cambridge University Press.

مقدمه‌ای قدیمی اما همچنان مفید برای مبانی آواشناسی آکوستیک.

Ladefoged, P. (1996) *Elements of Acoustic Phonetics*, 2nd edn., Chicago: University of Chicago Press.

مقدمه‌ای بسیار خواندنی برای آواشناسی آکوستیک، به‌ویژه در زمینه رابطه بین شکل‌موج آکوستیکی با بازنمایی طیفی آن.

Stevens, K. N. (1999) *Acoustic Phonetics*, Cambridge, MA: MIT Press.

یک مقدمه مهندسی معتبر برای نظریه آکوستیکی تولید گفتار.

تمرین

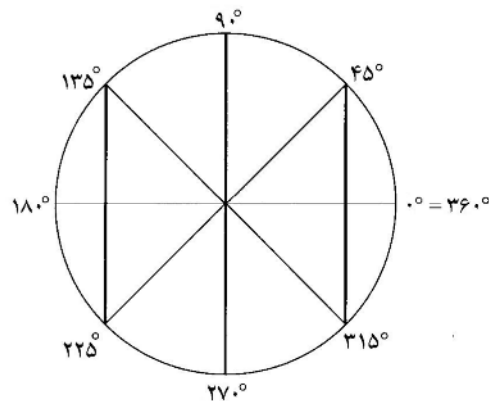
واژگان تخصصی

اصطلاحات زیر را تعریف کنید: صدا، رسانای آکوستیکی، شکل موج آکوستیکی، موج صوتی، رقت، تراکم، امواج متناوب، موج متناوب ساده، موج سینوسی، فرکانس، سیکل، دوره تناوب، هرتز، دامنه، فاز، موج متناوب مرکب، فرکانس پایه، مؤلفه‌های موج، طیف قدرت، قضیه فوریه، تجزیه فوریه، امواج نامتناوب، نوفه سفید، موج گذرا، پالس، صافی پایین‌گذر، نوار گذر، نوار دفع، صافی بالاگذر، شیب صافی، صافی نوارگذر، فرکانس مرکزی، پهنای باند.

پرسش

۱. چرا این جمله نادرست است: صدا را هنگامی احساس می‌کنید که ملکول‌های هوا از منبع صوتی به سوی پرده گوش حرکت می‌کنند.
۲. مقادیر زمانی روبه‌رو را برحسب ثانیه بیان کنید: ۱۰۰۰ میلی ثانیه، ۲۰۰ میلی ثانیه، ۱۰ میلی ثانیه، ۱۲۱۰ میلی ثانیه.
۳. اگر دوره تناوب یک موج سینوسی 0.01 ثانیه، 10 میلی ثانیه، 0.33 ثانیه و 33 میلی ثانیه باشد، فرکانس آن برحسب هرتز چقدر است؟
۴. یک موج سینوسی رسم کنید. برای این کار، ابتدا یک محور زمان رسم کنید و بر روی آن نقاطی را با فاصله مساوی 45 علامت‌گذاری نمایید. بدین صورت که اولین نقطه، صفر، نقطه بعدی (در سمت راست) 45 ، نقطه بعدی 90 و همین‌طور تا 720 ادامه دهید. همان‌طور که در شکل ۱-۱۵ می‌بینید، این نقاط درجه چرخش حول یک دایره را نشان می‌دهند (720 درجه، برابر با دو بار چرخش حول دایره است). سپس، مقادیر دامنه موج سینوسی را نیز به صورت ارتفاع خط‌های عمودی نسبت به خطی که از مرکز دایره از 0 تا 180 درجه می‌گذرد، علامت‌گذاری کنید. بدین ترتیب، مقدار دامنه در صفر درجه، مساوی

صفر است. مقدار آن در 45° درجه برابر با فاصله خط عمودی از خط مرکزی تا لبه دایره در 45° درجه است و به همین ترتیب، سایر مقادیر دامنه نیز به دست می‌آید. اگر خط‌های عمودی پایین خط مرکزی دایره قرار گیرند (مانند خط عمودی در 225° درجه) دامنه را به صورت مقداری منفی علامت‌گذاری کنید. حالا نقاط را به یکدیگر وصل نمایید. سعی کنید موج سینوسی را شبیه به شکل موج‌های این فصل رسم کنید.



شکل ۱-۱۵. درجات چرخش حول یک دایره

۵. موج مرکب حاصل از ترکیب امواج سینوسی 300 هرتزی و 500 هرتزی را (هر دو با قله دامنه ۱) رسم کنید.
۶. طیف موج متناوب مرکب حاصل از مؤلفه‌های 100 هرتزی و 700 هرتزی را (هر دو با قله دامنه ۱) رسم کنید.
۷. طیف صافی آکوستیکی حاصل از ترکیب دو صافی نوارگذر زیر را رسم کنید. فرکانس مرکزی صافی اول 500 هرتز و پهنای باند آن 50 هرتز است. فرکانس مرکزی صافی دوم 1500 هرتز و پهنای باند آن 150 هرتز است.

فصل ۲

نظریه آکوستیکی تولید گفتار: تولید شوا

ارتعاش تارهای صوتی، منبع رایج تولید صوت در واکه‌هاست و دستگاه گفتار، صافی آکوستیکی است که صوت ایجادشده توسط تارآواها را تغییر می‌دهد. این نظریه در آکوستیک گفتار به‌عنوان نظریه منبع-صافی تولید گفتار^۱ شناخته می‌شود (فانت^۲، ۱۹۶۰؛ فلانگان^۳، ۱۹۶۵). در این فصل، صوت حاصل از ارتعاش تارآواها را توصیف می‌کنیم و پس از معرفی اجمالی نظریه ذره‌ای^۴ استیونس (۱۹۷۲، ۱۹۸۹)، به بحث در مورد عملکرد صافی دستگاه گفتار در تولید شوا [ə] می‌پردازیم. نظریه آکوستیکی تولید گفتار، همچنان که از اسمش پیداست، نظریه‌ای فقط در مورد شوا نیست و مفاهیم معرفی‌شده در این فصل، در بخش دوم کتاب نیز مطرح می‌شوند.

۲-۱. واک

وقتی تارآواها به‌وسیله فشار زیر چاکنا به‌طور مکرر باز و توسط فشار کشسانی

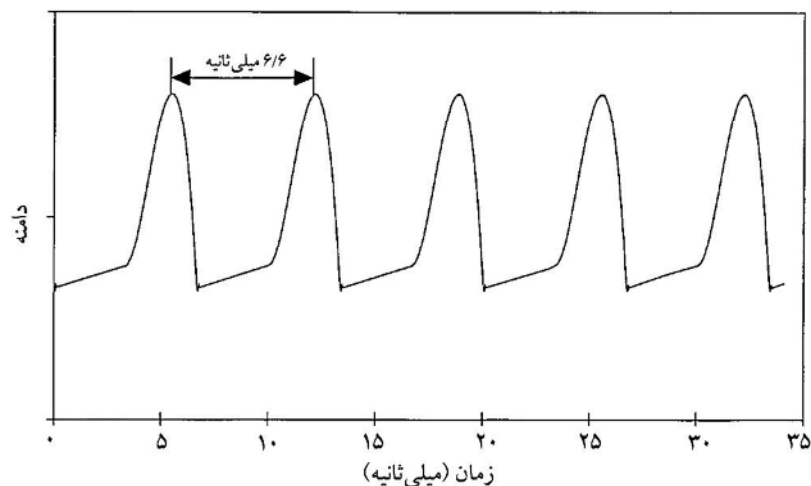
1. source-filter theory of speech production
3. Flanagan

2. Fant
4. quantal theory

ماهیچه وکالیس^۱ بسته می‌شوند، یک موج متناوب مرکب تولید می‌کنند. شکل ۱-۲ شکل موج واکِی را نشان می‌دهد که به وسیله یک سنتزکننده گفتار (با کیفیت بسیار طبیعی) تولید شده است. توجه داشته باشید که شکل موج واک یک الگوی تکرارشونده غیرسینوسی دارد. تعداد دفعاتی که این شکل موج متناوب مرکب در ثانیه تکرار می‌شود، فرکانس پایه (F_0) آن را تعیین می‌کند که با درک زیروبمی گفتار توسط شنونده مرتبط است. در این شکل، هر سیکل ۶۶ میلی‌ثانیه به طول می‌انجامد.

در نتیجه، فرکانس پایه ۱۵۰ هرتز است $(\frac{1 \text{ ثانیه}}{0.0066 \text{ ثانیه}})$

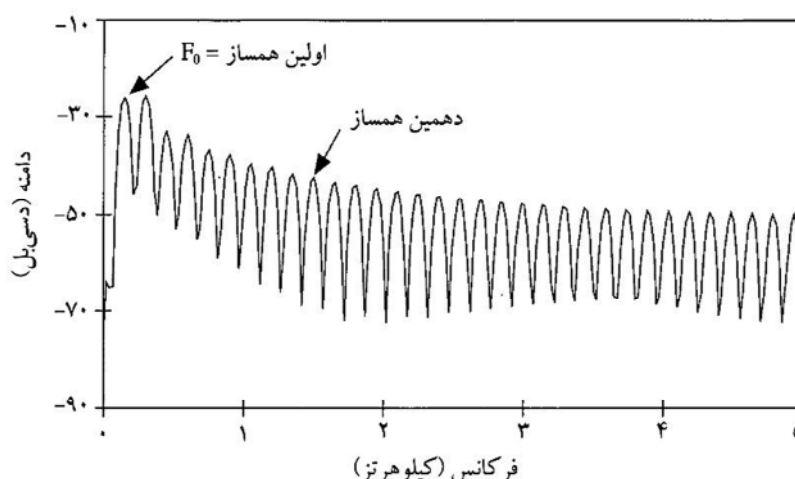
با تجزیه فوریه شکل موج واک، طیف قدرت آن که مؤلفه‌های فرکانسی و دامنه نوسان آنها را نشان می‌دهد (شکل ۲-۲)، به دست می‌آید. فرکانس پایه، اولین قله (با کمترین فرکانس) در طیف قدرت است و هر قله دیگری در طیف قدرت، مضربی از فرکانس پایه است.



شکل ۱-۲. پنج سیکل تولیدشده با منبع واک یک سنتزکننده گفتار (کلت و کلت^۲، ۱۹۹۰). در شبیه‌سازی آکوستیکی دستگاه گفتار این شکل موج به عنوان درون‌داد برای تولید واکه‌ها به کار می‌رود. فرکانس پایه ۱۵۰ هرتز است و هر سیکل ۶/۶ میلی‌ثانیه طول می‌کشد.

1. vocalis

2. Klatt and Klatt



شکل ۲-۲. طیف قدرت ارتعاش تارآواها که در شکل ۲-۱ نشان داده شد. همساز اول (که فرکانس آن برابر با فرکانس پایه واک است) در ۱۵۰ هرتز تولید می‌شود. بنابراین، همساز دهم در ۱۵۰۰ هرتز تولید می‌شود.

بنابراین، به عنوان مثال، قله دوم در طیف قدرت در ۳۰۰ هرتز، قله سوم در ۴۵۰ هرتز، و قله دهم در ۱۵۰۰ هرتز تولید می‌شود. مؤلفه‌های طیف قدرت واک، همساز^۱ نامیده می‌شوند. توجه داشته باشید که از همساز دوم به بعد دامنه هر همساز مقداری کمتر از همساز پیشین خود است. (دامنه نسبی همسازهای اول و دوم به نوع واک‌سازی^۲ بستگی دارد؛ در واک نفسی^۳ همساز اول بلندتر و همساز دوم نرم‌تر تولید می‌شود، در حالی که در واک بازداشته^۴ برعکس، همساز دوم بلندتر از اولی است. به شکل ۸-۳ رجوع کنید.)

در اینجا با ارائه مثالی نشان می‌دهیم که همسازها چه هستند. («جان اوها!»^۵ به این مثال اشکال وارد می‌کند، زیرا ارتعاش تارآواها به همان صورتی تولید صوت می‌کند که یک صوت با مدوله کردن^۶ جریانی از هوا می‌کند. به هر حال، روش بهتری برای نشان دادن اینکه همسازها، مدهای ارتعاش^۷ هستند، پیدا نکردم؛ هرچند که در مورد یک سیم مرتعش، مدهای مختلف ارتعاش در واقع فرکانس‌های بازخوانی سیم هستند. بنابراین، لازم به یادآوری است که این مثال، مثال خیلی خوبی نیست.) زمانی که شما یک سیم

1. harmonic

2. phonation type

3. breathy voicing

4. creaky voicing

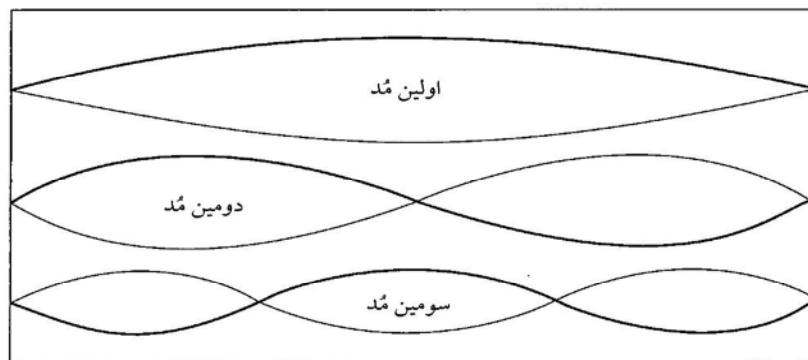
5. John Ohala

6. Modulating

7. modes of vibration

گیتار را به صدا درمی‌آورید، به ارتعاش درآمده و صوتی تولید می‌کند که تا حدودی شبیه واک است. با ارتعاش سیم گیتار یک صوت متناوب مرکب تولید می‌شود که دارای یک فرکانس پایه و همسازهایی با فرکانس بالاتر است که مضارب فرکانس پایه هستند. همسازهای نُت (به همراه فرکانس پایه) با روش‌های بیشماری که یک سیم گیتار می‌تواند ارتعاش کند، تولید می‌شوند. شکل ۲-۳ سه تا از این مُدهای ارتعاش را نشان می‌دهد. بالاترین نمودار در شکل، مُد دارای پایین‌ترین فرکانس ارتعاش است. کل سیم با فرکانسی که بر اساس طول سیم تعیین می‌شود، به سمت بالا و پایین نوسان می‌کند. این مُد ارتعاش، فرکانس پایه سیم گیتار و در نتیجه زیربمی نُت را مشخص می‌کند.

نمودار دوم در شکل، دومین مُد ارتعاش را نشان می‌دهد. مثل مُد اول (و در هر یک از مُدهای ارتعاش یک سیم گیتار) دو انتهای سیم حرکت نمی‌کند. اما در این مُد، طول سیم به دو قسمت تقسیم شده است. در حالی که نیمی از سیم به یک سمت حرکت می‌کند، نیمه دیگر آن به سمت مخالف در حرکت است و در وسط سیم، نقطه‌ای به نام گره^۱ وجود دارد که اصلاً حرکت نمی‌کند. از آنجا که فرکانس ارتعاش یک سیم به وسیله طول آن مشخص می‌شود و در دومین مُد ارتعاش، طول سیم نصف شده است، فرکانس دومین مُد دو برابر اولین مُد، یعنی دو برابر فرکانس پایه است.



شکل ۲-۳. سه مُد ارتعاش در سیمی که نواخته شده است. نمودار بالا الگوی حداکثر جابه‌جایی اولین مُد ارتعاش، نمودار وسط الگوی حداکثر جابه‌جایی دومین مُد ارتعاش، و نمودار پایین الگوی حداکثر جابه‌جایی سومین مُد ارتعاش را نشان می‌دهد.

پایین‌ترین نمودار در شکل ۲-۳ سومین مُد ارتعاش یک سیم گیتار را نشان می‌دهد. در این مُد، سیم به سه بخش تقسیم شده است. مثل مُد قبلی، دو انتهای سیم حرکت نمی‌کند (که «شرط مرزی» برای مُدهای ارتعاش است. با این مفهوم دوباره وقتی در مورد صافی دستگاه گفتار صحبت می‌کنیم، مواجه خواهیم شد). از آنجا که طول هر بخش ارتعاش‌کننده سیم، یک‌سوم طول کل سیم است، سومین مُد ارتعاش، منجر به ایجاد مؤلفه‌ای در موج مرکب (سومین همساز) می‌شود که سه برابر فرکانس پایه است.

یک راه برای شنیدن دومین همساز یک سیم گیتار وجود دارد. شما می‌توانید سیم را به صدا درآورید و سپس به آرامی وسط آن را لمس کنید. زمانی که این کار را انجام می‌دهید، صدا به دو صورت تغییر می‌کند: دامنه نوسان آن کاهش می‌یابد و زیربمی آن دو برابر می‌شود. زمانی که شما وسط سیم را (در نقاط حداکثر جابه‌جایی مُدهای ارتعاش) لمس می‌کنید، اولین و سومین همساز، میرا^۱ می‌شوند، در حالی که دومین همساز که حاصل دومین مُد ارتعاش است، میرا نمی‌شود. وقتی که شما وسط سیم را لمس می‌کنید، همه همسازهای فرد میرا می‌شوند، در نتیجه، تنها همسازهای زوج باقی می‌مانند.

مُدهای ارتعاش یک سیم گیتار، مثال ساده‌ای است از رابطه بین فرکانس پایه و همسازهای یک موج متناوب مرکب. ارتعاش تارآواها پیچیده‌تر از این است، اما روابط بین مؤلفه‌های شکل موج واک بیشتر به همین صورت است. ارتعاش تارآواها منجر به ایجاد یک صوت متناوب مرکب با همسازهایی می‌شود که مضارب صحیحی از فرکانس پایه هستند.

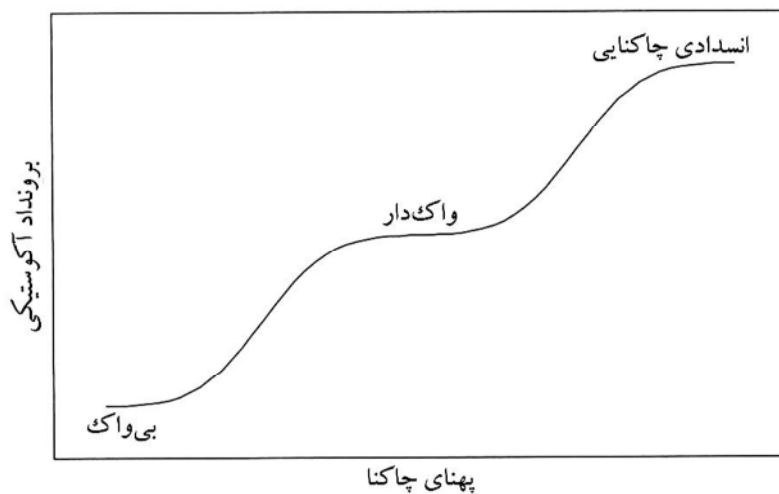
۲-۲. مقادیر ذره‌ای واک

عملکرد تارهای صوتی، یکی از روشن‌ترین مثال‌های نظریه ذره‌ای گفتار را به وجود می‌آورد. از این رو، در این بخش نگاهی به نظریه ذره‌ای می‌اندازیم. استیونس (۱۹۷۲) نظریه ذره‌ای را برای توجیه این واقعیت مطرح کرد که زبان‌ها از همه حالت‌های ممکن تولید گفتار استفاده نمی‌کنند. به‌ندرت در زبانی همه تولیدهای بیشمار ممکن برای ایجاد تمایز بین کلمات به کار رفته است. به نظر «استیونس» علت این است که نگاشت^۲ تولید به آکوستیک غیرخطی است. تولید تارآواها نمونه

1. damped

2. mapping

خوبی از این غیرخطی بودن است. طی گفتار، فضای بین تارآواها (چاکنا) از حالت کاملاً باز (زمانی که یک نفس عمیق می کشید) تا حالت کاملاً به هم چسبیده (هنگام تولید یک انسدادی چاکنایی) تغییر می کند. در شکل ۲-۴، محدوده پیوسته پهنای چاکنا بر روی محور افقی نشان داده شده است.



شکل ۲-۴. نمایش نواحی ذره‌ای به صورت فلات در رابطه بین پهنای چاکنا با برون داد آکوستیکی آن.

نظریه ذره‌ای و پیچ رادیو

نظریه ذره‌ای بر اساس این ایده بنا شده است که نگاشت غیرخطی تولید به آکوستیک، تعیین کننده مشخصه‌های تمایز دهنده موجود در زبان است. اکنون این سؤال پیش می آید که «منظور از غیرخطی بودن چیست؟»

کنترل خطی شبیه پیچ تنظیم صدای رادیو است. اگر پیچ را بچرخانید، صدا بلندتر (یا یواش تر) می شود. هر قدر پیچ را بیشتر بچرخانید، تغییر بیشتری در بلندی صدا ایجاد می شود. کنترل غیرخطی شبیه پیچ تنظیم موج رادیو است. با چرخاندن این پیچ، رادیو ایستگاه‌های متفاوتی را دریافت می کند. بنابراین، یک چرخش کوچک پیچ ممکن است منجر به تغییر زیادی در صدای رادیو شود، اما یک چرخش بزرگ ممکن است تغییر خیلی کمی (فقط پارازیت) ایجاد کند.

در شکل، پیامد آکوستیکی تغییر تدریجی پهنای چاکنا از باز تا بسته (با فرض مقدار ثابت فشار هوای پشت تارهای صوتی) بر روی محور عمودی نشان داده شده است. بُعد تولیدی در این مثال، پهنای چاکنا، مفهومی انتزاعی است، به این معنی که شامل هم ناحیه چاکناست و هم شامل کشش ماهیچه‌هایی که چاکنا و خود تارآواها را می‌بندند. در ابتدا، وقتی تارآواها باز هستند، سایش [h] تولید می‌شود، اما به تدریج با کم شدن پهنای چاکنا و رسیدن به میزانی از پهنای آن، تغییری ذره‌ای در برون‌داد آکوستیکی رخ می‌دهد و تارهای صوتی شروع به ارتعاش می‌کنند. ارتعاش ممکن است تا حدودی نفسی باشد (مثل تولید [h] در ahead)، اما نوع واک تولیدشده کاملاً با قبل متفاوت است. وقتی که چاکنا بسته می‌شود، تغییر ناگهانی دیگری در برون‌داد آکوستیکی روی می‌دهد که نتیجه آن تولید یک انسدادی چاکنایی است. بنابراین، پهنای چاکنا مانند پیچ تنظیم موج رادیو است و نه پیچ بلندی صدای رادیو؛ بدین معنی که یک تغییر کوچک می‌تواند تأثیر زیادی بر جای بگذارد.

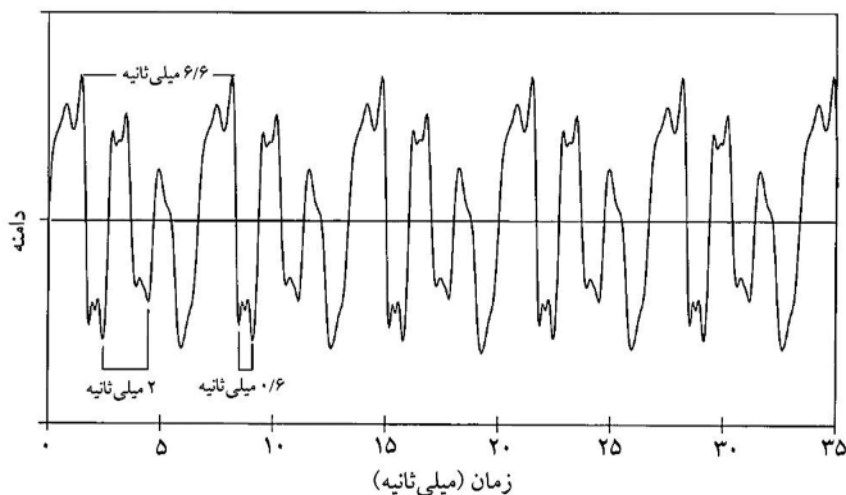
«استیونس» معتقد است که زبان‌ها از چنین روابط غیرخطی بین تولید و آکوستیک استفاده می‌کنند. نواحی افقی در نمودار شکل ۲-۴ با عنوان «بی‌واک»، «واک‌دار»، و «انسداد چاکنایی»، نواحی پایدار^۱ در نگاشت تولید به آکوستیک محسوب می‌شوند. به عنوان مثال، گویشوران می‌توانند پهنای مختلف از چاکنا را انتخاب کرده و در عین حال، همچنان واک تولید کنند. از آنجا که برون‌داد آکوستیکی در پهنای متفاوتی از چاکنا عملاً یکسان است، مقداری لغزش تولیدی^۲ امکان‌پذیر است. به این ترتیب، غیرخطی بودن طبیعی در نگاشت تولید به آکوستیک منجر به ایجاد طبقات طبیعی آواهای گفتاری می‌شود. در دیدگاه استیونس، با تعیین فهرست روابط غیرخطی در نگاشت تولید به آکوستیک، فهرست مشخصه‌های تمایزدهنده آوایی که برای تمایز معنایی در زبان به کار می‌رود، به دست می‌آید. تعیین نواحی پایدار (نواحی ذره‌ای) در مورد فعالیت تارهای صوتی آسان است، اما موارد دیگری را خواهیم دید که بحث در آنها اندکی از این پیچیده‌تر است.

1. regions of stability

2. articulatory slop

۳-۲. صافی دستگاه گفتار

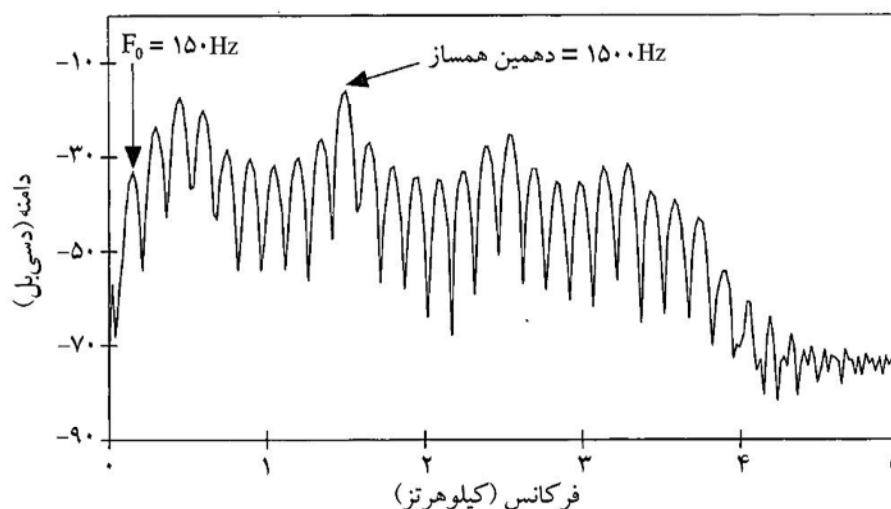
دستگاه گفتار می‌تواند به‌عنوان یک صافی عمل کرده و شکل موج واکی را که در شکل ۱-۲ نشان داده شد، را به یک موج متناوب مرکب مانند شکل ۲-۵ تبدیل کند. این شکل با عبور واک ستنز شده شکل ۱-۲ از مجموعه‌ای از صافی‌های نوارگذر رسم شده است. واک ستنز شده حاصل، شبیه شوآ [ə]، واک هجای اول کلمه *ahead* در انگلیسی آمریکایی است. دوره تناوب آن همچنان $6/6$ میلی‌ثانیه است، بنابراین الگوی شکل موج 150 بار در ثانیه تکرار می‌شود. اما در این شکل موج، مؤلفه‌های نوسانی مهم دیگری هم قابل تشخیص است. به‌عنوان مثال، در هر دوره تناوب موج، مؤلفه‌ای وجود دارد که هر 2 میلی‌ثانیه یک بار تکرار می‌شود و در نتیجه فرکانس آن 500 هرتز است. همچنین، توجه کنید که یک نوسان حتی سریع‌تر در شکل موج وجود دارد که به‌صورت یک الگوی دوقله‌ای سوار بر برخی قله‌های 2 میلی‌ثانیه‌ای دیده می‌شود. این نوسان هر $0/666$ میلی‌ثانیه (فاصله بین قله‌های دوتایی) یک بار تکرار می‌شود و بنابراین، فرکانس آن 1500 هرتز است.



شکل ۲-۵. شکل موج آکوستیکی یک شوآی ستنز شده. دوره تناوب سه مؤلفه اصلی آن در شکل مشخص شده است.

طیف قدرت منبع واک (شکل ۲-۲) و طیف قدرت شوآ در شکل ۲-۶، فرکانس پایه یکسانی دارند (همان طور که همسازهای ۱۵۰ هرتزی، ۳۰۰ هرتزی، ۴۵۰ هرتزی و غیره نشان می دهند). بنابراین، هر دو موج از همسازهای یکسانی تشکیل شده اند، اما دامنه نسبی همسازها متفاوت است. در طیف قدرت منبع واک، با افزایش فرکانس، دامنه همسازها به طور یکنواختی کاهش پیدا می کند؛ در حالی که در طیف شوآ، همسازها در حدود ۵۰۰، ۱۵۰۰، ۲۵۰۰ و ۳۵۰۰ هرتز، دامنه نوسان بالاتری نسبت به سایر همسازها در فرکانس های دیگر دارند. همان طور که شکل ۲-۵ نشان می دهد، این تغییرات به ظاهر کوچک در دامنه نسبی همسازها، تأثیر بسزایی بر صورت شکل موج دارد.

در این آوا (شوآی سنتز شده)، قله های پهن طیفی که شواهد آنها را هم در شکل موج و هم در طیف داریم، با عبور شکل موج واک از مجموعه ای از فیلترهای نوارگذر با فرکانس های مرکزی ۵۰۰، ۱۵۰۰، ۲۵۰۰ و ۳۵۰۰ هرتز تولید شده اند. این واقعیت که با چنین روشی در سنتز گفتار، می توان یک واکه قابل فهم تولید کرد، نشان می دهد که دستگاه گفتار مانند مجموعه ای از صافی های نوارگذر عمل می کند.



شکل ۲-۶. طیف قدرت واکه ای که در شکل ۲-۵ نمایش داده شد. توجه کنید که همسازها همان فرکانس هایی را دارند که در طیف قدرت واک داشتند (شکل ۲-۲). به عنوان مثال، فرکانس دهمین همساز ۱۵۰۰ هرتز است. اما دامنه نسبی همسازها یکسان نیست. در طیف قدرت شوآ، قله های پهن طیفی در حدود ۵۰۰، ۱۵۰۰، ۲۵۰۰ و ۳۵۰۰ هرتز مشاهده می شود.

۲-۴. پاندول، امواج ایستاده^۱ و سازه‌های واکه^۲

از آنجا که دستگاه گفتار دارای فرکانس‌های بازخوانی^۳ است، می‌توان ویژگی‌های آکوستیکی شوآ را با استفاده از بانکی از صافی‌های نوارگذر شبیه‌سازی کرد. همسازهایی از منبع واک که تقریباً همان فرکانس‌هایی را دارند که بازخوان‌های دستگاه گفتار دارند، تقویت می‌شوند (گویی که در نوارگذر یک صافی قرار گرفته‌اند) و مؤلفه‌هایی از منبع واک که فرکانس آنها حول‌وحوش یک بازخوان نیست، میرا می‌شوند.

فرکانس بازخوانی یک پاندول

فرکانس‌های بازخوانی شوآ به وسیلهٔ امواج ایستاده در دستگاه گفتار به وجود می‌آیند. نوسان یک پاندول بعضی از ویژگی‌های بازخوانی را نشان می‌دهد که در اینجا نگاهی به آنها می‌اندازیم. برای این کار آزمایش زیر را انجام دهید. وزنه‌ای را به یک سیم وصل کنید و ببینید چگونه نوسان می‌کند (می‌توانید از یک بند کفش و یک قاشق استفاده کنید). اگر انتهای سیم را در یک دستتان نگه دارید و با دست دیگر وزنه را به یک طرف بکشید و سپس رها کنید، برای مدتی وزنه به جلو و عقب نوسان می‌کند. توجه کنید که مدت‌زمان نوسان وزنه در یک سیکل رفت‌وبرگشت، مستقل از بزرگی نوسان است. شما می‌توانید وزنه را با حرکت دادن میچ دست نیز به نوسان درآورید. اگر میچ دستتان را در یک زمان‌بندی درست از سویی به سوی دیگر حرکت دهید، پاندول در قوس‌های بزرگ نوسان می‌کند. توجه کنید که اگر زمان‌بندی درست باشد، اندازهٔ قوس‌ها خیلی بزرگ‌تر از حرکت میچ دست شماست. اگر شما میچ‌تان را خیلی سریع یا خیلی کند حرکت دهید، وزنه به جای اینکه نوسان کند، در یک‌جا می‌ایستد.

چرا مهم است که پاندول را چقدر آهسته یا چقدر سریع به نوسان درآورید؟ مدت‌زمانی که یک نوسان، کامل می‌شود، ثابت است. هر سیکل مدت‌زمان یکسانی

1. standing waves

2. vowel formants

3. resonant frequency

برای کامل شدن صرف می‌کند، چه تازه با حرکات کوچک مچ دست شروع به نوسان دادن پاندول کرده باشید یا با حرکات بزرگ بازو آن را به شدت به نوسان درآورید، و چه در حالی که انتهای سیم را ثابت نگه داشته‌اید به وزنه با یک دست فقط ضربه بزنید. این نشان می‌دهد که پاندول، یک فرکانس بازخوانی مشخص دارد. اگر طول سیم را کوتاه کنید، طول دوره تناوب کاهش می‌یابد. می‌توانید اهمیت طول سیم را با مقایسه فرکانس‌های بازخوانی پاندول‌هایی با وزنه‌های مختلف نیز اثبات کنید (به عنوان مثال، یک قاشق را با یک چکش مقایسه کنید). بنابراین، پاندول یک فرکانس بازخوانی طبیعی دارد. اگر به سیستم در محدوده فرکانس بازخوانی انرژی وارد کنید (با حرکت مچ دست)، آنگاه سیستم با افزودن هر ذره انرژی متوالی که به آن وارد می‌شود، انرژی را تقویت می‌کند (حرکات مچ دست، اگر درست زمان‌بندی شوند، به یکدیگر اضافه می‌شوند)، اما اگر در فرکانس‌های دیگری به سیستم انرژی وارد کنید، سیستم انرژی را ضعیف می‌کند، زیرا انرژی به وسیله پالس‌های انرژی قبلی خنثی می‌شود.

بنابراین، هر حرکت مچ دست مقدار کمی انرژی به پاندول وارد می‌کند. زمان‌بندی حرکات مچ دست، فرکانسی است که در آن انرژی به سیستم وارد می‌شود و نوسان پاندول (یا فقدان آن)، نتیجه جفت‌شدگی^۱ بین منبع انرژی و سیستم بازخوانی است.

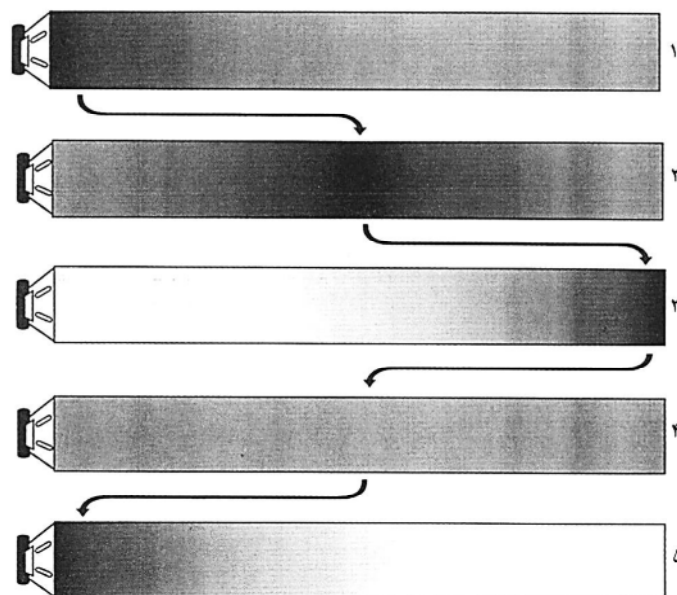
مثال آکوستیکی بازخوان به صورت زیر است. فرض کنید یک لوله برداشته و یک بلندگو در یک سر آن و یک درپوش در سر دیگر آن می‌گذارید و یک موج سینوسی از طریق بلندگو وارد آن می‌کنید. شکل ۲-۷ پاسخ لوله آکوستیکی را هنگامی که موج سینوسی شروع به بیرون آمدن از بلندگو می‌کند، نشان می‌دهد. در این شکل، فشار هوا با سایه مشخص شده است؛ به این صورت که فشار اتمسفری، خاکستری است، تراکم، سیاه و رقت، سفید. سایه‌های شکل یادآور مهمی هستند از اینکه شکل موج آکوستیکی گفتار یک موج طولی است (جابه‌جایی ذرات بر روی

1. coupling

همان محوری است که موج حرکت می‌کند) و نه یک موج عرضی (جابه‌جایی ذرات عمود بر جهت موج). برای ما بسیار ساده‌تر است که نوسانات فشار سینوسی را به صورت موج‌های سینوسی که بر روی صفحه بالا و پایین می‌روند، نشان دهیم، اما باید در خاطر داشت که صدا به صورت یک موج طولی حرکت می‌کند. می‌توانیم از یک اسباب‌بازی فنری بنام «اسلینکی»^۱ برای نشان دادن امواج طولی و عرضی استفاده کنیم. اگر یک انتهای اسلینکی را در دست نگه داشته و بگذارید انتهای آزاد آن به سمت زمین آویزان شود، می‌توانید با حرکت دادن دستتان به جلو و عقب به صورت افقی، یک موج عرضی در آن بفرستید. با حرکت دادن دست به سمت بالا و پایین می‌توانید یک موج طولی در آن بفرستید. حلقه‌های فنر در موج طولی، به هم نزدیک و از هم دور می‌شوند، تراکم و رقت!

یک قله تراکم در زمان ۱ درست از دهانه بلندگو حرکت کرده و در زمان ۳ به انتهای لوله می‌رسد (این موج «موج انتشار» نامیده می‌شود). این قله فشار در زمان ۴ از انتهای درپوش‌دار لوله منعکس می‌شود («موج انعکاسی») در حالی که دره فشاری که در زمان ۳ نزدیک بلندگو پدید آمده بود، به وسط لوله گسترش یافته است. این دو قله و دره یکدیگر را خشی می‌کنند و در نتیجه، فشار در تمام نقاط لوله معادل فشار اتمسفری می‌شود. در زمان ۵، قله فشار مثبت به بلندگو برمی‌گردد، در حالی که دره فشار در انتهای دیگر لوله به درپوش می‌رسد. حال اگر موج سینوسی که از بلندگو بیرون می‌آید، فرکانس مناسبی داشته باشد، یک موج تراکم جدید به انعکاس موج اول اضافه می‌شود. (در واقع، فرکانس موج سینوسی برای زمان‌بندی دره فشار در زمان ۳ نیز حیاتی است.) از این رو، در زمان ۵، در انتهای بلندگودار لوله، یک قله فشار وجود دارد که ترکیبی است از قله منعکس‌شده و قله جدیدی که از بلندگو پخش شده است، مثل هل دادن کودک بر روی تاب. با هر سیکل موج سینوسی، قله فشار دامنه بالاتری پیدا می‌کند و دره فشار پایین‌تر می‌رود. من در واقع، چنین لوله‌ای را نساخته‌ام، اما اگر پس از چند سیکل منفجر شود، تعجب‌آور نیست.

1. Slinky

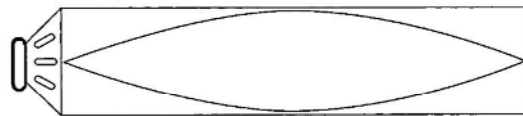


شکل ۲-۷. اولین سیکل یک موج سینوسی نواخته شده در یک لوله آکوستیکی که هر دو انتهای آن مسدود است (یک انتهای آن با یک بلندگو بسته شده است). نوارهای شماره ۱ تا ۵ فشار هوای داخل لوله را در پنج زمان مختلف نشان می دهند. مقدار فشار با درجه تیرگی نشان داده شده است، به این صورت که سیاه تراکم، سفید رقت و خاکستری فشار صفر (اتمسفری) را نشان می دهد.

در زمان های بعدی، شکل نوسان به صورت حالت های زمان ۳ و ۵ با حالت زمان ۴ در بین آنهاست. شکل ۲-۸ (الف) آن را نشان می دهد. این شکل، اولین بازخوان یک لوله آکوستیکی را نشان می دهد که مانند لوله مذکور هر دو انتهای آن بسته است. زمان های ۱ تا ۵ در شکل، متناظر با نوارهای سایه دار شکل ۲-۷ است. الگوی نوسان در لوله «ساکن می ماند»، به این معنا که در نقطه وسط لوله، فشار همواره صفر است، اما در دو انتهای لوله، فشار بین قله ها و دره ها نوسان می کند.

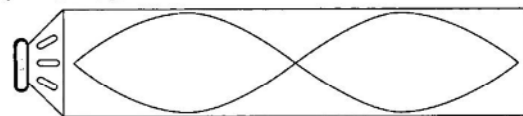
برای بازنمایی موج های سینوسی در شکل ۲-۸ از برخی علائم استفاده شد. بدین صورت که علامت + نشان دهنده قله های فشار موج و صفر نشانگر نقاطی است که موج در آنها دامنه صفر را قطع می کند. علامت - برای نشان دادن دره های فشار و علامت $\pm$ نشان دهنده فشار صفری است که وقتی یک قله و یک دره یکدیگر را خنثی می کنند، ایجاد می شود.

اولین بازخوان (الف)



۱	+		
۲	۰	+	
۳	-	۰	+
۴	۰	±	۰
۵	+	۰	-
۶	۰	±	۰
۷	-	۰	+
۸	۰	±	۰
۹	+	۰	-

دومین بازخوان (ب)

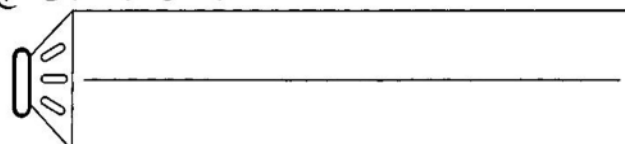


۱	+		
۲	-	+	
۳	+	-	+
۴	-	+	-
۵	+	-	+
۶	-	+	-
۷	+	-	+
۸	-	+	-
۹	+	-	+

(ادامه دارد)

شکل ۲-۸ موج‌های ایستاده دو پایین‌ترین فرکانس‌های بازخوانی در لوله‌ای که در هر دو انتها مسدود است، به صورتی که در یک انتهای آن یک بلندگو قرار دارد. زمان به صورت عمودی از ابتدا (۱) تا انتها (۹) نشان داده شده است. قله‌های فشار با مثبت (+)، دره‌های فشار با منفی (-)، و صفر یا فشار اتمسفری، با صفر (۰)، و در مورد قله‌ها و دره‌هایی که یکدیگر را خنثی می‌کنند، فشار با مثبت-منفی (±) بازنمایی شده است. شکل (الف)، اولین بازخوان لوله، متناظر با موج شکل ۲-۷ را نشان می‌دهد. شکل (ب) دومین فرکانس بازخوانی لوله و شکل (ج) الگوی تولیدشده به وسیله یک فرکانس غیربازخوانی را نشان می‌دهند. موج‌های سینوسی داخل لوله، امواج ایستاده را نمایش می‌دهند. در این نوع بازنمایی، گره‌ها به صورت نقاطی که در آنها امواج سینوسی یکدیگر را قطع می‌کنند (کمترین جابه‌جایی ذرات) و پادگره‌ها به صورت نقاطی که در آنها امواج سینوسی بیشترین فاصله را از یکدیگر دارند (حداکثر جابه‌جایی ذرات) نشان داده شده‌اند.

یک فرکانس غیربازخوانی (ج)



۱	+				
۱/۵		+			
۲			+		
۲/۵	-			+	
۳		-			+
۳/۵			-	+	
۴	+		+	-	
۴/۵		+			-
۵	+		+	-	
۵/۵	-	+	-	+	
۶		-	+		+
۶/۵	-		-	+	
۷	+	-	+	-	+
۷/۵		+	-	+	-
۸	+		+	-	
۸/۵	-	+	-	+	-

اگر به ستون مقادیر فشار در انتهای بلندگودار لوله در شکل ۲-۸ (الف) نگاه کنید، نوسان بین قله‌ها و دره‌ها در موج سینوسی را که از بلندگو پخش می‌شود، می‌بینید. علاوه بر این، اگر از زمان ۴ شروع کنیم، به‌طور منظم یک الگوی فشار تکرارشونده در لوله وجود دارد. در زمان‌های فرد، یک قله فشار در یک انتهای لوله و یک دره فشار در انتهای دیگر آن وجود دارد، در حالی که در زمان‌های زوج، قله و دره در وسط لوله یکدیگر را خنثی می‌کنند. بنابراین، در وسط لوله، فشار مساوی صفر، معادل فشار اتمسفری باقی می‌ماند، در حالی که در دو سر لوله، فشار بین قله‌های مثبت و منفی

نوسان می‌کند. این الگو به عنوان یک موج ایستاده شناخته می‌شود. موجی که در لوله شکل ۸-۲ رسم شده، بازنمایی استاندارد یک موج ایستاده است. این بازنمایی، میزان جابه‌جایی ذرات را به عنوان تابعی از طول لوله نشان می‌دهد. به عنوان مثال، ما وسط لوله را به عنوان نقطه‌ای مشخص کردیم که فشار هوای لوله در آنجا همیشه صفر است، زیرا در این نقطه موج انتشار و موج انعکاسی یکدیگر را ختشی می‌کنند. جالب این است که چون فشار در این نقطه صفر باقی می‌ماند، در حالی که امواج صوتی از آنجا می‌گذرند، جابه‌جایی هر ذره نسبتاً زیاد است و در نتیجه، میانگین فاصله بین ذرات (فشار هوا) ثابت باقی می‌ماند. بنابراین، نقطه وسط لوله، نقطه حداکثر جابه‌جایی ذرات است. به همین صورت، دیدیم که در هر دو انتهای لوله نیز نوسانات فشار نسبتاً زیاد است، اما انتهای بسته لوله در این نقاط، مانع جابه‌جایی ذرات می‌شود. در نتیجه، در دو انتهای لوله، نقاطی با حداقل جابه‌جایی ذرات داریم. در موج ایستاده، نقطه وسط لوله که در آن فشار صفر باقی می‌ماند، پادگره^۱ نامیده می‌شود، در حالی که دو نقطه انتهای لوله که در آنها فشار در حداکثر مثبت یا منفی است، گره نامیده می‌شود. امواج سینوسی که در لوله شکل ۸-۲ (الف) رسم شده‌اند، با توزیع جابه‌جایی ذرات مطابقت دارند. در این شکل، گره‌ها به صورت نقاط تلاقی امواج سینوسی و پادگره‌ها به شکل قله‌های امواج نشان داده شده‌اند.

برخلاف پاندول که یک سیستم هماهنگ ساده خوانده می‌شود به این دلیل که تنها یک فرکانس بازخوانی دارد، لوله‌های آکوستیکی فرکانس‌های بازخوانی متعددی دارند. بازنمایی‌های (الف) و (ب) در شکل ۸-۲، پایین‌ترین فرکانس‌های بازخوانی لوله‌ای را نمایش می‌دهند که در هر دو انتها بسته است. به علاوه، شکل (ب) پاسخ لوله به یک موج سینوسی را نشان می‌دهد که فرکانسی دو برابر فرکانس موج شکل (الف) دارد. بنابراین، با نگاهی به اولین ستون علائم مثبت و منفی می‌توانید ببینید که در حالی که در بازخوان اول (شکل الف)، دره موج سینوسی در زمان ۳ رخ می‌دهد، در بازخوان دوم

1. antinode

(شکل ب)، یک سیکل موج سینوسی در این زمان کامل شده است. لازم به ذکر است که قله فشار منتشرشده به وسیله بلندگو در هر دو بازخوان با سرعت یکسانی حرکت می‌کند. در زمان ۳، قله به انتهای درپوش‌دار لوله می‌رسد و در زمان ۵ به بلندگو گردد. مدت زمانی که طول می‌کشد تا موج صوتی به انتهای درپوش‌دار لوله برسد و سپس بازگردد، تعیین‌کننده فرکانس‌هایی است که در لوله تشدید یا بازخوان می‌شوند. برای هر دو بازخوان اول و دوم لازم است که موج سینوسی تولیدشده به وسیله بلندگو در زمان ۵ یک قله فشار داشته باشد. اگر طول لوله بلندتر باشد و در نتیجه، مدت‌زمان لازم برای رفتن موج به انتهای درپوش‌دار لوله و سپس برگشت آن، طولانی‌تر شود، در این صورت، لوله تنها موج‌های سینوسی را بازخوان می‌کند که دوره تناوب طولانی‌تری (فرکانس پایین‌تر) داشته باشند. از این رو، یک قله فشار منتشرشده با بلندگو به‌عنوان مثال در زمان ۶ رخ می‌دهد، نه در زمان ۵.

حال ببینیم اگر فرکانس موج سینوسی منتشرشده به وسیله بلندگو با فرکانس بازخوانی لوله مطابقت نداشته باشد، چه اتفاقی رخ می‌دهد. شکل ۲-۸ (ج) مثالی از آن را نشان می‌دهد. آنچه در این شکل می‌بینیم، پیامد انتشار یک موج سینوسی به وسیله بلندگوست که فرکانس آن مابین فرکانس موج‌های نشان‌داده‌شده در شکل‌های (الف) و (ب) است. سیکل این موج با چهار واحد زمانی کامل می‌شود (در این شکل برخلاف شکل‌های قبل، زمان به صورت نصف واحدهای زمانی قبلی نشان داده شده است). در شکل (ج) قله انعکاسی در چندین نقطه با قله جدیدی که به وسیله بلندگو منتشر شده، ترکیب می‌شود. در همین نقاط، دره فشار انعکاسی نیز با دره منتشرشده ترکیب می‌شود. اما این دو همیشه در جهت مخالف حرکت می‌کنند، به این معنا که هرگز به هم اضافه نمی‌شوند تا موج ایستاده‌ای با دامنه‌ای بزرگتر از دامنه موجی که بلندگو منتشر می‌کند، بسازند. این مثال نشان می‌دهد که برای اینکه موج صوتی ورودی تشکیل یک موج ایستاده در لوله بدهد، بایستی موج سینوسی تولیدشده به وسیله بلندگو، قله‌ای در زمان ۵ داشته باشد (زمانی که در آن موقع موج انعکاسی به بلندگو بازگشته است).

طول موج

اگر می‌خواهید بدانید که آیا یک موج سینوسی در یک لوله بازخوان می‌شود یا نه، لازم است فاصله‌ای را که در یک سیکل می‌پیماید، بدانید. این فاصله طول موج نامیده می‌شود. اگر فرض کنید که موج نشان داده شده در شکل ۷-۲ از انتهای لوله منعکس نشده و در یک لوله بلند به حرکت خود ادامه می‌دهد، می‌توانید تصویری از طول موج پیدا کنید. برای محاسبه طول موج، به دیرش^۱ دوره تناوب موج سینوسی نیاز داریم. گفتیم که بازه زمانی بین دو قله متوالی موج، دوره تناوب (T) نامیده می‌شود (به عنوان مثال، بازه بین زمان‌های ۱ و ۵ در شکل ۷-۲). طول موج (λ) به سرعت صوت در لوله نیز بستگی دارد. (ممکن است با مسئله به گونه دیگر آشنا باشید: اگر با سرعت ۳۰ مایل در ساعت در حرکت باشید و با این سرعت به مدت یک ساعت سفر کنید، چه مسافتی را می‌پیمایید؟) در هوای گرم و مرطوب، سرعت صوت (c) حدود ۳۵۰۰۰ سانتی‌متر در ثانیه (cm/s) است. بنابراین، طول موج برابر است با دیرش دوره تناوب ضربدر سرعت صوت ($\lambda = cT$)، یا به عبارت دیگر، از آنجا که فرکانس، معکوس دوره تناوب است ($1/T = f$)، طول موج مساوی است با سرعت صوت تقسیم بر فرکانس موج سینوسی ($\lambda = c/f$). به عنوان مثال، دیرش دوره تناوب یک موج سینوسی ۱۰۰۰ هرتزی، ۱ ثانیه یا ۰/۰۰۱ ثانیه است. در نتیجه، طول موج آن ۳۵ سانتی‌متر است ($35000 \times 0.001 = 35000/1000 = 35$). یعنی قله‌های متوالی در یک موج سینوسی ۱۰۰۰ هرتزی، ۳۵ سانتی‌متر از هم فاصله دارند. از این رو، اگر دو میکروفن به فاصله ۳۵ سانتی‌متر از یکدیگر قرار داده و یک موج سینوسی ۱۰۰۰ هرتزی نزدیک آنها تولید کنید، میکروفن‌ها در هر لحظه از زمان، نقاط یکسانی را در فاز موج دریافت می‌کنند، در حالی که اگر فاصله میکروفن‌ها از هم ۱۷/۵ سانتی‌متر باشد، در لحظه‌ای که یکی از آنها یک قله ضبط می‌کند، دیگری یک دره ضبط می‌کند. هر دو میکروفن یک موج سینوسی ضبط می‌کنند، اما فاز نسبی آنها به دلیل طول موج به فاصله بین آنها بستگی دارد.

1. duration

بنابراین، در مثالی که ملاحظه کردیم (لوله‌ای که هر دو سرش بسته است)، آن امواج سینوسی در لوله تولید موج ایستاده می‌کنند که قالب^۱ لوله باشند. به عبارت دیگر، امواج سینوسی که قالب لوله هستند به شکلی در لوله جلو و عقب می‌روند که درست در همان زمانی که قله یا دره فشاری به وسیله بلندگو تولید می‌شود، یک قله تراکم یا یک دره رقت مشابه در انتهای بلندگودار لوله دارند. اگر طول لوله و سرعت صوت در هوا را بدانید، می‌توانید فرکانس‌های تولیدکننده امواج ایستاده را محاسبه کنید (یعنی فرکانس‌های بازخوانی لوله را). طول موج اولین بازخوان (λ) دو برابر طول لوله $(2L)$ است. حرکت از بلندگو تا درپوش، نیمی از طول موج را دربرمی‌گیرد و مسیر برگشت نیمه دیگر آن. از این رو، فرکانس بازخوانی عبارت است از سرعت صوت تقسیم بر طول موج $(f=c/2L)$. به عنوان مثال، اگر طول لوله‌ای ۸ سانتی‌متر باشد، آن‌گاه طول موج بازخوان اول 16 سانتی‌متر است و بنابراین، فرکانس آن، $2187/5 \text{ Hz} = 35000/16$ است. سایر بازخوان‌های بعدی لوله، مضاربی از اولین فرکانس بازخوانی هستند. مثلاً، طول موج بازخوان دوم (شکل ۲-۸ (ب) را ببینید) نصف طول موج بازخوان اول است و بنابراین، دومین فرکانس بازخوانی یک لوله ۸ سانتی‌متری مساوی $f_2=c/L$ است. به طور کلی فرکانس‌های بازخوانی لوله‌ای را که در هر دو انتها بسته است، می‌توان با معادله (۱-۲) محاسبه کرد. در این معادله، n شماره بازخوان، c سرعت صوت و L طول لوله است. بنابراین، فرکانس‌های بازخوانی یک لوله ۸ سانتی‌متری که در هر دو انتها بسته است، برابر با $2187/5$ هرتز، 4375 هرتز، $6562/5$ هرتز و غیره است.

$$f_n = nc/2L \quad (1-2)$$

حال فرض کنید که از بلندگو یک موج مرکب پخش می‌شود و نه یک موج سینوسی. همان‌طور که در فصل ۱ دیدیم، موج مرکب ترکیبی از تعدادی مؤلفه‌های سینوسی است. از این رو، تنها تفاوت بین مثال‌های شکل ۲-۸ و یک موج مرکب در این است که امواج سینوسی بیشمار هم‌زمان وارد لوله می‌شوند. بنابراین، به جای در نظر گرفتن ویژگی‌های بازخوانی لوله بر اساس پاسخ لوله به هر موج سینوسی به

1. fit

نوبت، ممکن است ترکیبی از همه امواج نشان داده شده در شکل ۲-۸ و بیشتر را داشته باشیم. تصور نمی شود راهی برای توصیف آن بر اساس امواجی که در لوله بالا و پایین می دهند، وجود داشته باشد، زیرا وضعیت بسیار پیچیده است. اما می توان پیش بینی هایی ارائه نمود. پیش بینی نخست این است که اگر منبع صوتی مرکب، دارای فرکانس هایی نزدیک به فرکانس های بازخوانی لوله باشد (مانند دو مثالی که در شکل های ۲-۸ (الف و ب) نشان داده شد)، دامنه این مؤلفه های منبع تقویت می شود. پیش بینی دوم این است که اگر منبع صوتی مرکب دارای فرکانس هایی باشد که به فرکانس های بازخوانی لوله نزدیک نباشند، در این صورت، دامنه این مؤلفه ها تضعیف می شود. از این رو، اگرچه در این بخش مبحث موج ایستاده در رابطه با امواج سینوسی درون یک لوله مطرح شد، اما این مبحث با تحلیل آکوستیکی گفتار مرتبط است زیرا موج مرکب را می توان به مجموعه ای از امواج سینوسی تجزیه کرد.

روش دیگری هم برای توصیف امواج ایستاده که در بالا گفته شد، وجود دارد که با تولید طبیعی گفتار متفاوت است. برای روشن شدن مبحث، لوله ای را فرض کنید که در هر دو انتها بسته است. واضح است که در شوا و بیشتر آواهای گفتاری دیگر، علاقمند به آکوستیک لوله هایی هستیم که در یک انتها باز هستند. تنها تفاوت آکوستیکی بین لوله ای که در هر دو انتها بسته است با لوله ای که فقط در یک انتها بسته است، مربوط به این است که چگونه امواج صوتی از انتهای باز لوله منعکس می شوند. ممکن است درک اینکه صوت از انتهای باز لوله منعکس می شود، دشوار باشد، اما این اتفاق روی می دهد، زیرا موج صوتی به حجم ساکن و بزرگی از هوا که درست بیرون لبه لوله قرار دارد، برخورد می کند. شما ممکن است بی آنکه متوجه شده باشید، این نوع انعکاس صوت را تجربه کرده باشید. به عنوان مثال، وقتی فردی در اتومبیل به طرف پنجره خم شده و سعی می کند از طریق پنجره باز اتومبیل با شما صحبت کند، پنجره مانند انتهای باز لوله عمل کرده و بیشتر انرژی آکوستیکی موجود در سیگنال گفتاری به جای آنکه به بیرون برود، به داخل ماشین منعکس می شود. از این رو، هرچند که آن شخص ممکن است واضح صحبت کند، اما به سختی می توانید گفته هایش را بشنوید.

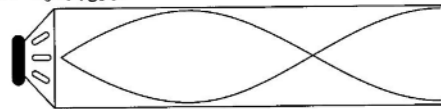
گفتیم که امواج صوتی از انتهای باز لوله به درون آن منعکس می‌شوند. اما به هر حال، نوع انعکاسی که در انتهای باز لوله صورت می‌گیرد با نوعی که در انتهای بسته صورت می‌گیرد، تفاوت دارد. تفاوت در این است که علامت موج صوتی تغییر می‌کند، به نحوی که یک قله به صورت یک دره از انتهای باز لوله منعکس می‌شود. مثالی از این نوع تغییر قطبیت^۱ را می‌توان در صدای شَرَقْ شَرَقْی که شلاق (نوار بلندی از چرم بافته شده با یک دسته کوتاه) ایجاد می‌کند، ملاحظه نمود. برای اینکه شلاق، صدای شَرَقْ شَرَقْ تولید کند، موجی در طول آن بفرستید. وقتی که موج به انتهای شلاق می‌رسد که شبیه به انتهای باز یک لوله است (زیرا به هیچ چیز متصل نیست)، دستخوش یک تغییر قطبیت ناگهانی می‌شود و یک صدای بلند تولید می‌کند! تغییر قطبیتی که در انعکاس امواج صوتی در انتهای باز یک لوله روی می‌دهد، تولید صدای بلند نمی‌کند، اما فرکانس‌های بازخوانی لوله را از اساس تغییر می‌دهد.

شکل ۲-۹ مثل شکل ۲-۸ دو پایین‌ترین فرکانس بازخوانی یک لوله را که در یک انتها باز است، نشان می‌دهد. ابتدا به دومین فرکانس بازخوانی نگاه کنید (شکل الف). با تعقیب قله تراکمی که به وسیله بلندگو در زمان ۱ منتشر شده، می‌بینیم که در زمان ۴ به انتهای لوله می‌رسد و دچار یک تغییر قطبیت از + به - می‌شود (در این شکل برای روشن شدن مطلب، مقیاس زمانی متفاوتی در مقایسه با شکل ۲-۸ به کار رفته است). یکی از پیامدهای تغییر قطبیت در این موج ایستاده این است که فشار در انتهای باز لوله همواره صفر است، زیرا قله در این نقطه دره می‌شود و این دو همدیگر را خنثی می‌کنند. به دنبال تغییر قطبیت در انتهای باز لوله، قله فشار منتشر شده به وسیله بلندگو در زمان ۱، به صورت یک دره فشار در زمان ۷ به بلندگو بازمی‌گردد. سپس، دوباره به سمت انتهای باز لوله حرکت می‌کند و دوباره دچار تغییر قطبیت شده و به بلندگو در زمان ۱۳ بازمی‌گردد. از این رو، دوره تناوب پایین‌ترین فرکانس بازخوانی لوله ۱۲ است. این، در شکل ۲-۹ (ب) که با یک مقیاس زمانی فشرده نشان داده شده، قابل مشاهده است. موج سینوسی منتشر شده به وسیله بلندگو برای آنکه با قله یا دره فشار

1. polarity change

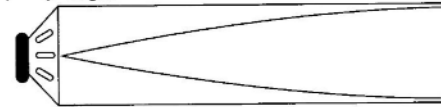
منعکس شده از انتهای باز لوله هماهنگ شود، بایستی فرکانس نشان داده شده در ستون اول شکل (ب) را داشته باشد. موج انعکاسی بایستی ۴ بار طول لوله را بپیماید قبل از اینکه به عنوان یک قله فشار به بلندگو بازگردد. این، یعنی اینکه طول موج اولین فرکانس بازخوانی، چهار برابر طول لوله است (4L).

دومین بازخوان (الف)



۱	+			
۲		+		
۳	-		+	
۴		-		±
۵	+		-	
۶		±		±
۷	-		-	
۸		±		±
۹	+		-	
۱۰		±		±
۱۱	-		+	
۱۲		±		±
۱۳	+		-	

اولین بازخوان (ب)



۱	+			±
۴				
۷	-			±
۱۰				
۱۳	+			±
۱۶				
۱۹	-			±
۲۲				
۲۵	+			±
۲۸				
۳۱	-			±
۳۴				
۳۷	+			±

شکل ۲-۹. امواج ایستاده دو پایین ترین فرکانس بازخوانی لوله ای که در یک انتها با یک بلندگو مسدود و در انتهای دیگر باز است. زمان به صورت عمودی از ابتدا تا انتها نشان داده شده است. قله های فشار با مثبت (+)، دره های فشار با منفی (-)، و فشار اتمسفری با صفر، و در مورد قله ها و دره هایی که یکدیگر را خنثی می کنند، با مثبت-منفی (±) نشان داده شده است. شکل (الف) دومین بازخوان و شکل (ب) اولین بازخوان لوله را نشان می دهد.

اصلاح طول^۱

اصلاح طول در آواشناسی آکوستیک پرداختن به جزئیات خسته‌کننده است. طول مؤثر لوله‌ای که در یک انتها باز است، در واقع کمی بلندتر از طول واقعی آن است. بنابراین، قطر انتهایی باز لوله d را در عدد 0.4 ضرب کنید و حاصل آن را به طول لوله اضافه نمایید. بدین ترتیب، این واقعیت که **انعکاس آکوستیکی** در انتهای باز لوله کمی «نرم» است و نقطه انعکاس در واقع، کمی فراتر از انتهای لوله است، جبران می‌شود. بنابراین، فرمول محاسبه فرکانس‌های بازخوانی لوله‌ای که در یک انتها باز است، مانند دستگاه گفتار، به صورت زیر می‌باشد:

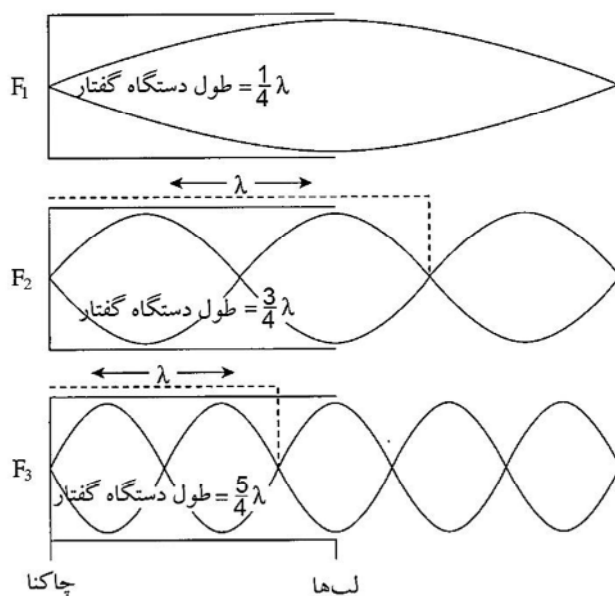
$$F_n = (2n - 1)c/4(L + 0.4d)$$

شکل ۲-۹ (الف)، دومین فرکانس بازخوانی را نشان می‌دهد. این شکل نشان می‌دهد که در این نوع لوله رابطه بین فرکانس‌های بازخوانی به سادگی لوله‌ای که در هر دو انتها بسته است، نیست. اگر به یاد داشته باشید، گفتیم که در لوله‌ای که در هر دو انتها مسدود است، فرکانس‌های بازخوانی در مضارب پایین‌ترین بازخوان تولید می‌شوند؛ یعنی دومین فرکانس بازخوانی دو برابر اولین فرکانس بازخوانی است و به همین ترتیب، الی آخر. اما اینجا در لوله‌ای که در یک انتها باز است، برعکس، دیرش دوره تناوب دومین فرکانس بازخوانی، یک‌سوم دیرش دوره تناوب اولین بازخوان است و در نتیجه، دومین فرکانس بازخوانی سه برابر اولین فرکانس بازخوانی است. الگوی کلی آن به این صورت است که فرکانس‌های بازخوانی لوله‌ای که در یک انتها بسته و در انتهای دیگر باز است، مضارب فرد اولین بازخوان هستند: یعنی دومین بازخوان سه برابر اولین بازخوان است و سومین بازخوان پنج برابر اولین بازخوان و به همین ترتیب. می‌توان رابطه بین شماره بازخوان و مضارب فرد را به صورت $(2n-1)$ بیان کرد. در این رابطه، n شماره فرکانس بازخوانی است. دومین

1. end correction

بازخوان سه برابر اولین بازخوان (۱-۴)، سومین بازخوان پنج برابر اولین بازخوان (۱-۶) و به همین صورت الی آخر.

بر اساس آنچه گفته شد، می توان ویژگی های آکوستیکی دستگاه گفتار در تولید شوآ را به صورت لوله ای که در یک انتها بسته (چاکنا) و در یک انتها باز (لب ها) است، مدل سازی نمود. بازخوان های طبیعی چنین لوله ای با چند شرط مرزی شبیه شرایط مرزی همسازهای یک سیم مواجه هستند. شرایط مرزی فرکانس های بازخوانی لوله (که در نمودارهای جابه جایی ذرات موج ایستاده در شکل ۲-۹ نشان داده شده اند) ناشی از این واقعیت است که امواج صوتی، نوسانات فشاری هستند که در فضا حرکت می کنند و بنابراین، فاصله ای که یک سیکل موج می پیماید تا کامل شود (طول موج آن)، برحسب اینچ یا میلی متر قابل اندازه گیری است. در امواج ایستاده نقاطی در فضا وجود دارد که در آنها جابه جایی ذرات در حداکثر است و نقاط دیگری که جابه جایی ذرات در آنها در حداقل است. طول موج بازخوان های لوله ای که در یک انتها بسته و در انتهای دیگر باز است، به نحوی است که نقاط دارای حداکثر میزان جابه جایی ذرات (پادگره ها) در امواج ایستاده در انتهای باز لوله روی می دهند، در حالی که نقاطی که جابه جایی ذرات در آنها صفر است (گره ها)، در انتهای بسته رخ می دهند. این شرایط مرزی برای فرکانس های بازخوانی بر اساس سازگاری موج صوتی با لوله تعیین می شود. همان طور که ارتعاش یک سیم گیتار به دلیل اینکه هر دو انتهای سیم ثابت است، محدود است (شکل ۲-۳)، به همین ترتیب، بازخوان های دستگاه گفتار نیز به دلیل ماهیت انعکاس آکوستیکی در هر یک از دو انتهای دستگاه گفتار، محدود می شوند. انتهای بسته لوله با فشار بالا سازگار است و انتهای باز لوله با فشار صفر. شکل ۲-۱۰ طول موج نسبی سه بازخوان اول دستگاه گفتار را که «سازه های» واکه نامیده می شوند، در شوآ نشان می دهد. این امواج صوتی به نحوی با دستگاه گفتار هماهنگ می شوند که یک حداکثر فشار در چاکنا و یک حداقل فشار در لب ها وجود داشته باشد.



شکل ۲-۱۰. سه پایین ترین موج فشاری که شرایط مرزی برای بازخوان شدن در یک دستگاه گفتار فاقد انسداد را دارند، شبیه وضعیت دستگاه گفتار هنگام تولید واژه [ə]. این سه سازه اول به صورت F_1 - F_3 نام گذاری می شوند.

در شکل ۲-۱۰ امواج ایستاده به صورت امواج سینوسی در یک لوله بازنمایی شده اند. طول موج پایین ترین فرکانس بازخوانی که در شکل با F_1 نشان داده شده، چهار برابر طول لوله است. در این موج ایستاده نیز مانند همه بازخوان های دیگر دستگاه گفتار، حداقل جابه جایی ذرات در چاکنا و حداکثر جابه جایی ذرات در لبها روی می دهد. می توانید تصور کنید که امواج شکل ۲-۱۰ را مثل فنر می کشید یا به هم می فشارید. طول موج های نشان داده شده در شکل از بین همه طول موج های ممکن انتخاب شده اند، زیرا شرایط مرزی برای بازخوان شدن در دستگاه گفتار را دارند. از آنجا که فرکانس موج صوتی برابر است با سرعت صوت تقسیم بر طول موج ($f=c/\lambda$)، و از آنجا که طول موج بازخوان های لوله ای که در یک انتها باز و در انتهای دیگر بسته است، در صورتی که طول لوله را بدانید، قابل محاسبه است، بنابراین، اگر طول لوله را بدانیم، می توانیم فرکانس های بازخوانی یک لوله واحد را

که در یک انتها بسته و در انتهای دیگر باز است، محاسبه کنیم. با استفاده از معادله‌های زیر (۲-۲)، سه بازخوان اول بر اساس طول لوله به دست می‌آید.

$$F_1 = c/\lambda_1 = c/(4L) = c/4L$$

$$F_2 = c/\lambda_2 = c/(4/3L) = 3c/4L \quad (2-2)$$

$$F_3 = c/\lambda_3 = c/(4/5L) = 5c/4L$$

اولین عبارت در هر یک از این فرکانس‌های بازخوانی مربوط به طول موج ایستاده در دستگاه گفتار است. دومین عبارت، طول موج موج ایستاده را برحسب طول دستگاه گفتار نشان می‌دهد و آخرین عبارت، ساده‌سازی جبری عبارت دوم است (هر بازخوان طول موجی دارد که مضرب فردی از سرعت صوت تقسیم بر چهار برابر طول دستگاه گفتار است). می‌توان عبارت پایانی فرکانس‌های بازخوانی لوله‌ای را که در یک انتها باز و در انتهای دیگر بسته است، به صورت معادله (۲-۳) خلاصه کرد. در این رابطه n شماره بازخوان (سازه)، c سرعت صوت در هوای گرم و مرطوب (35000 cm/s)، و L طول لوله برحسب سانتی‌متر است (با فرض افزودن یک اصلاح طول کوچک $0.4d$ به آن، به عبارت دیگر، محاسبه طول «مؤثر» لوله).

$$F_n = \frac{(2n-1)c}{4L} \quad (2-3)$$

طبق این فرمول، پایین‌ترین فرکانس بازخوانی در دستگاه گفتار یک مرد با طول مؤثر $17/5$ سانتی‌متر برابر با 35000 cm/s تقسیم بر چهار ضربدر $17/5$ سانتی‌متر، یعنی 500 هرتز است. اگر دستگاه گفتار کوتاه‌تر باشد (مثلاً 15 سانتی‌متر)، پایین‌ترین فرکانس بازخوانی بالاتر است (583 هرتز). دومین بازخوان دستگاه گفتار با ضرب سرعت صوت در 3 و سپس تقسیم آن بر $4L$ به دست می‌آید. بنابراین، در یک دستگاه گفتار $17/5$ سانتی‌متری، دومین بازخوان 1500 هرتز و در یک دستگاه گفتار 15 سانتی‌متری، 1750 هرتز است. فرکانس‌های بازخوانی دستگاه گفتار را سازه نیز می‌نامند که با شروع از پایین‌ترین فرکانس به صورت F_1, F_2, F_3 و غیره نشان داده می‌شوند.

لوله‌های مستقیم و بازخوان‌های عرضی

مدل‌سازی دستگاه گفتار به صورت یک لوله مستقیم با یک موج صوتی که در طول آن حرکت می‌کند، سؤال برانگیز است. اگر لوله انحنای داشته باشد، چه وضعیتی پیش می‌آید (همان‌طور که دستگاه گفتار دارد)؟ و چرا نباید امواج صوتی که در عرض لوله بازخوان می‌شوند، علاوه بر آنهایی که در طول لوله بازخوان می‌شوند، نیز در نظر گرفت؟ جواب دو سؤال فوق «هیچ» و «البته، چرا که نه» است.

امواج صوتی توانایی حرکت در خمیدگی‌ها را دارند. به عنوان مثال، وقتی در یک شلنگ آب حرف می‌زنید، اگر شلنگ را خم کنید، صدا تغییری نمی‌کند. البته از آنجا که شلنگ یک لوله بلند است، مانند یک صافی آکوستیکی عمل کرده و صدای شما را تغییر می‌دهد. اما با خم کردن شلنگ تغییر بیشتری در سیگنال روی نمی‌دهد. نوسانات فشاری که موج صوتی هستند، در همه جهات از منبع صوت پخش می‌شوند. دیواره‌های لوله، مانع این الگوی انتشار می‌شوند، اما اگر لوله خم شود، گرایش طبیعی موج صوتی برای انتشار یکسان در همه جهات، سبب می‌شود که موج در انحنای حرکت کند.

آیا پهنای لوله نیز موجب تولید بازخوان‌هایی شبیه بازخوان‌های طول لوله که به آنها پرداختیم، می‌شود؟ پاسخ «مثبت» است اما فرکانس این «بازخوان‌های عرضی» معمولاً آنقدر بالاست که هیچ تأثیر واقعی بر ویژگی‌های آکوستیکی گفتار (به ویژه واژه‌ها) ندارد. به عنوان مثال، طول موج یک بازخوان را در نظر بگیرید که در لوله شکل ۲-۷ بالا و پایین می‌رود (نه جلو و عقب). طول موج پایین‌ترین بازخوان بسیار کوتاه (دو برابر عرض لوله) است و در نتیجه، فرکانس بسیار بالایی دارد. این بازخوان‌های عرضی در آواهای گفتاری وجود دارند، اما فرکانس آنها به قدری بالاست که به دلایل عملی قابل چشم‌پوشی هستند.

یک بار دیگر به طیف قدرت شکل ۲-۶ نگاه کنید. این طیف، برون‌داد آکوستیکی دستگاه گفتاری را نشان می‌دهد که انسداد ندارد (قطر یکسان)، طول آن ۱۷/۵ سانتی‌متر است و ارتعاش تارهای صوتی با فرکانس پایه ۱۵۰ هرتز روی می‌دهد. بازخوان‌های دستگاه گفتار سبب تولید قله‌های طیف در ۵۰۰ هرتز، ۱۵۰۰ هرتز، ۲۵۰۰ هرتز، و ۳۵۰۰ هرتز می‌شوند. همان‌گونه که طیف نشان می‌دهد، بازخوان‌های دستگاه گفتار، همسازهایی را که فرکانس نزدیک به فرکانس‌های بازخوانی دارند، تقویت و سایر همسازها را میرا می‌کنند. همسازهایی که فرکانس آنها با فرکانس‌های بازخوانی مطابقت ندارد، همچنان در سیگنال وجود دارند، اما به دلیل فیلتر شدن به وسیله دستگاه گفتار، نسبتاً ضعیف‌تر هستند.

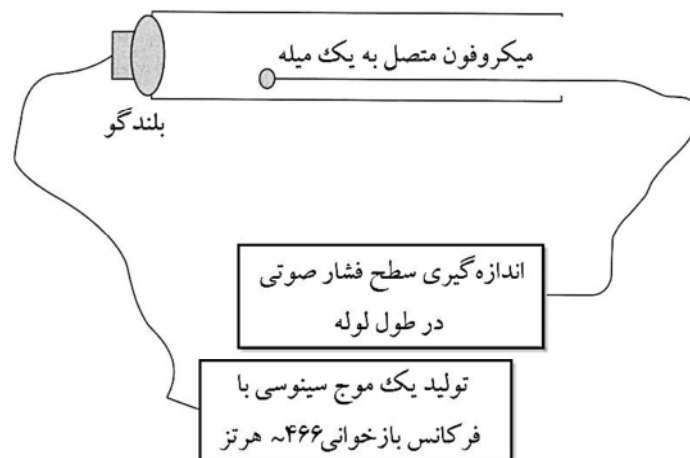
در این فصل عملکرد صافی دستگاه گفتار را وقتی که هیچ گرفتگی در آن وجود ندارد، در تولید واکه شوآ [ə] مورد بررسی قرار دادیم. همان‌طور که دیدیم، فرکانس‌های بازخوانی دستگاه گفتار (سازه‌ها) بر اساس طول دستگاه گفتار تعیین می‌شوند. در نتیجه، در شرایط مساوی، افرادی که دستگاه گفتار کوتاهی دارند، دارای سازه‌های بالاتری نسبت به افرادی که دستگاه گفتار بلندتری دارند، هستند. به علاوه، از آنجا که پیش‌آمدگی لب‌ها و پایین‌رفتن حنجره نیز طول دستگاه گفتار را افزایش می‌دهد. در نتیجه، فرکانس همه سازه‌ها به دلیل این تغییرات پایین‌تر است.

۲-۵. جستجوی گره‌ها و پادگره‌ها در یک لوله آکوستیکی

وقتی در اینترنت برای مشاهده فرکانس‌های بازخوانی و امواج ایستاده به جستجو می‌پردازید، با فیلم‌ها و برنامه‌های معرفی مواجه می‌شوید. من به‌ویژه ویدیویی از فیلم *لوله روبن*^۱ را دوست دارم که با آهنگی از «دیو بروبک»^۲ برای کنترل شعله آتش پایان می‌یابد.

۱. Ruben's tube. وسیله‌ای برای نشان دادن امواج ایستاده در لوله با استفاده از شعله آتش. (مترجم)

2. Dave Brubeck

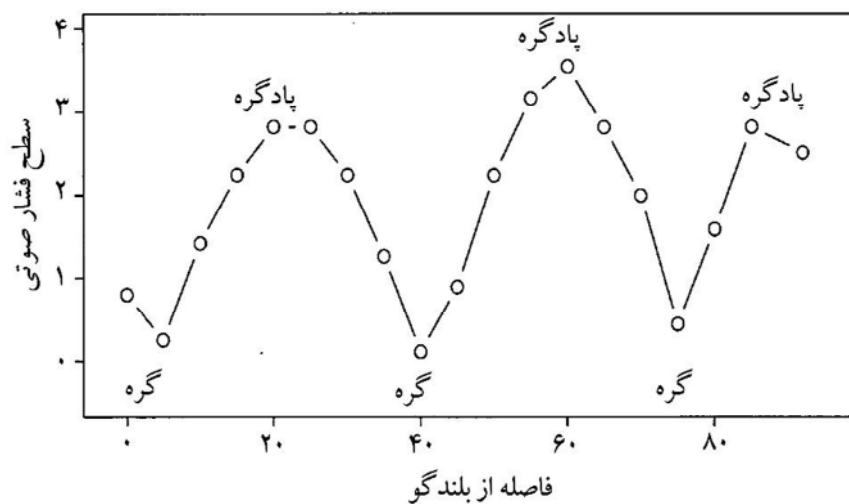


شکل ۲-۱۱. شکل کلی وسیله به کاررفته برای یافتن گره‌ها و پادگره‌ها در یک لوله آکوستیکی بازخوانی. خوب، در کسل‌کننده‌ترین آزمایشی که همانا عنوان این بخش است، از وسیله نشان داده شده در شکل ۲-۱۱ استفاده شد. یک لوله پلاستیکی شفاف داریم که طول آن ۹۲ سانتی‌متر و قطر آن در حدود ۳ سانتی‌متر است. از یک طرف یک بلندگوی کوچک به انتهای لوله کاملاً کیپ نصب می‌شود. انتهای دیگر لوله باز است و یک «میکروفون متصل به یک میله» داخل آن قرار می‌دهیم (میکروفون دکمه‌ای کوچکی که روی یک میله فلزی باریک و بلند به طول ۱۰۰ سانتی‌متر چسبانده شده است). بلندگو به یک ابزار مولد موج سینوسی وصل است که برای انتشار یک موج سینوسی ۴۶۶ هرتزی تنظیم شده است (مقدار F_3 لوله: $466 = (93/2 \times 4) / (5 \times 343 \times 0.01)$ ، با فرض سرعت صوت در دمای خشک اتاق و با اصلاح طولی که طول مؤثر لوله را به $93/2$ سانتی‌متر می‌رساند). سپس دامنه موج سینوسی را در فواصل تقریباً مساوی در امتداد طول لوله، با لغزاندن میکروفون از انتهای باز لوله به انتهای بسته آن اندازه می‌گیریم.

نتیجه جستجوی ساده موج ایستاده در لوله‌ای که در یک انتها بسته و در انتهای دیگر باز است، در شکل ۲-۱۲ نشان داده شده است. محور افقی، محل میکروفون را در لوله نشان می‌دهد، صفر یعنی میکروفون در انتهای بسته لوله است. محور عمودی، دامنه

نوسان موج صوتی را نشان می‌دهد. الگویی که در این شکل می‌بینیم، بسیار شبیه نمودار F3 موج ایستاده در شکل ۲-۱۰ است. در نقاطی که در موج ایستاده به عنوان «گره» مشخص شده است، یعنی نقاط دارای حداقل جابه‌جایی ذرات، مقدار دامنه تقریباً صفر اندازه‌گیری شد، اما در نقاطی که در موج ایستاده به عنوان «پادگره» مشخص شده، یعنی نقاط دارای حداکثر جابه‌جایی ذرات، مقدار دامنه موج بالاست. این جالب نیست که حتی وقتی میکروفن کاملاً نزدیک به بلندگو قرار گرفته، اندازه دامنه صفر نیست و مقداری دامنه قابل اندازه‌گیری است؟ امواج ایستاده واقعاً در آنجا هستند!

اگر ویدیویی از لوله «روبن» در اینترنت ببینید، متوجه خواهید شد که شعله‌ها در پادگره‌ها بالاتر از گره‌ها هستند (این متداول‌ترین رفتار لوله «روبن» است). لوله روبن نمایشی واقعاً جالب از امواج ایستاده نیز هست. علت اینکه شعله‌ها در پادگره‌ها بالاتر هستند به این دلیل است که فشار هوا در پادگره‌ها ثابت است، در حالی که در گره‌ها فشار هوا نوسان می‌کند. مثال میکروفن روی میله، به ما نشان داد که میکروفن جابه‌جایی ذرات را می‌گیرد (که تبدیل به جابه‌جایی غشای میکروفن می‌شود).



شکل ۲-۱۲. محل حداکثر سطح فشار صوتی (پادگره‌ها: نقاط حداکثر جابه‌جایی ذرات) و حداقل سطح فشار صوتی (گره‌ها: نقاط حداقل جابه‌جایی ذرات) در سومین فرکانس بازخوانی لوله شکل ۲-۱۱.

منابعی برای مطالعه بیشتر

- Fant, C. (1960) *Acoustic Theory of Speech Production*, The Hague: Mouton.
- اثری برجسته در آواشناسی آکوستیک، با ارائه استنتاجی ریاضی برای ویژگی‌های آکوستیکی بیشتر آواهای گفتاری.
- Flanagan, J. L. (1965) *Speech Analysis Synthesis and Perception*, Berlin: Springer-Verlag.
- مطالعه‌ای بسیار گسترده در زمینه علم و فناوری گفتار به همراه نتایج اولیه نظریه آکوستیکی تولید گفتار.
- Klatt, D. H. and Klatt, Laura (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820-57. A description of the KLSYN90 speech synthesis program.
- این برنامه، ابزاری ارزشمند به‌ویژه برای دانشجویان آواشناسی آکوستیک است، زیرا پیاده‌سازی مستقیم نظریه آکوستیکی تولید گفتار است و از سوی بسیاری از آزمایشگاه‌های پژوهشی در جهان برای استفاده آزادانه در دسترس است.
- Stevens, K. N. (1972) The quantal nature of speech: Evidence from articulatory-acoustic data. In E. E. David, Jr. and P. B. Denes (eds.), *Human Communication: A unified view*, New York: McGraw-Hill, 51-66.
- نخستین انتشار نظریه ذره‌ای استیونس.
- Stevens, K. N. (1989) On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.
- جدیدترین گزارش به‌روزشده نظریه ذره‌ای.

تمرین

واژگان تخصصی

اصطلاحات زیر را تعریف کنید: واک، فرکانس پایه واک، همساز، غیرخطی، ناحیه ذره‌ای، فرکانس‌های بازخوانی، موج ایستاده، پادگره، گره، سیستم هارمونیک ساده، طول موج، سرعت صوت (به cm/s)، تغییر قطبیت، انعکاس آکوستیکی، سازه، F_1 ، F_2 ، F_3 ، لوله «روبن».

پرسش

۱. در طیف شکل ۲-۲ مقدار کجی طیفی^۱ را به دست آورید. یک خط مستقیم (تا حد امکان نزدیک) بر قله همسازها مماس کنید. هیچ خط مستقیمی همه نقاط را دقیقاً به هم وصل نمی‌کند، اما سعی کنید مجموع فاصله مطلق بین قله‌ها و خط تا حد امکان کم شود (که به آن «رگرسیون خطی» می‌گویند). اکنون مقدار دامنه خط را در ۱ کیلوهرتز و ۲ کیلوهرتز اندازه بگیرید. کجی طیفی، همان شیب خط است که به صورت میزان تغییرات در ۱ کیلوهرتز بیان می‌شود.
۲. سه فرکانس بازخوانی اول لوله‌ای را که در هر دو انتها بسته است، برای هر یک از طول‌های ۴ سانتی‌متر، ۱۰ سانتی‌متر و ۱۲ سانتی‌متر محاسبه کنید.
۳. سه فرکانس بازخوانی اول لوله‌ای را که در یک انتها بسته و در انتهای دیگر باز است، برای هر یک از طول‌های ۱۲ سانتی‌متر، ۱۵ سانتی‌متر و ۱۸ سانتی‌متر محاسبه کنید. مقدار فرکانس‌های بازخوانی هنگامی که یک اصلاح طول برای لوله‌ای با قطر ۳ سانتی‌متر اضافه شود، چقدر است؟
۴. فرض کنید که گردی لب‌ها، ۱ سانتی‌متر به طول هر یک از لوله‌های سؤال ۳ می‌افزاید. آیا این افزایش، تغییر فرکانسی یکسانی (برحسب هرتز) در هر یک از سازه‌ها ایجاد می‌کند؟ آیا این افزایش، تغییر فرکانسی یکسانی در F_2 (برحسب هرتز) بر حسب هر یک از طول‌های دستگاه گفتار ایجاد می‌کند؟

1. spectral tilt

فصل ۳

پردازش دیجیتالی سیگنال^۱

در دهه‌های اخیر، ابزار مورد استفاده در آواشناسی آکوستیک به‌طرز چشمگیری تغییر کرده است. امروزه به‌جای اسیلوسکوپ، دستگاه ثبت موج^۲ و طیف‌نگار^۳، تقریباً فقط از کامپیوتر و برنامه‌های کامپیوتری استفاده می‌کنیم. در این فصل، مبحث «پردازش دیجیتالی سیگنال» یعنی روش به‌کارگیری سیگنال آکوستیکی به‌وسیله کامپیوترها مطرح می‌شود. هدف این است که خواننده درک عمیقی از اینکه کامپیوترها چگونه تجزیه آکوستیکی را انجام می‌دهند، پیدا کند. مباحث از جهتی جالب نیز هستند، زیرا به کلمات و اصطلاحاتی اشاره می‌شود که ممکن است در رابطه با ابزارهای پخش موسیقی و یا سیستم‌های صوتی کامپیوتر شنیده باشید. (برای مطالب بیشتر در زمینه مباحث این فصل، کتاب ریچارد لاینز^۴ ۱۹۹۷ بسیار توصیه می‌شود).

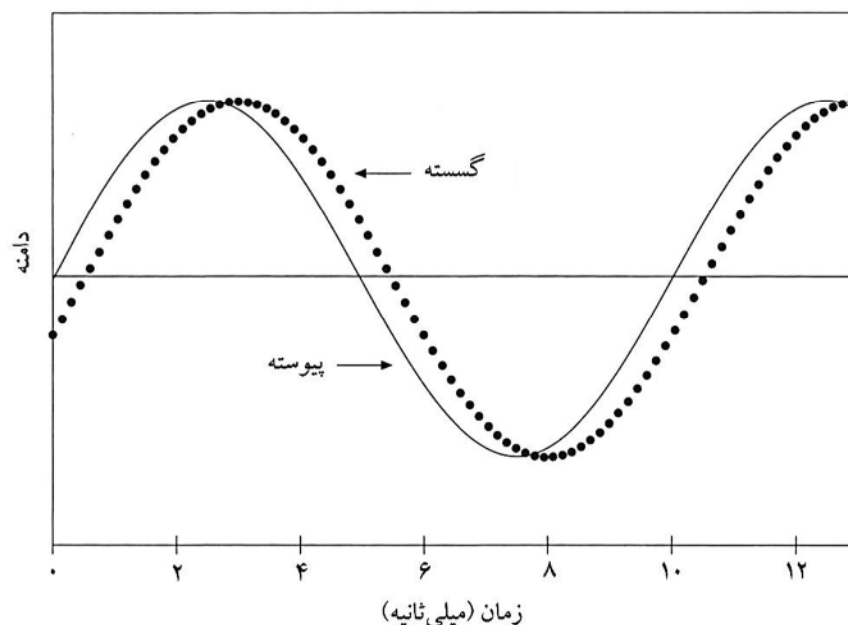
۳-۱. سیگنال‌های پیوسته در مقابل سیگنال‌های گسسته

درک فرایند پردازش دیجیتالی سیگنال (دی‌اس‌پی) از تفاوت بین سیگنال‌های

1. digital signal processing (DSP)
3. spectrograph

2. chart recorder
4. Richard Lyons

پیوسته و گسسته آغاز می‌شود. این تفاوت مربوط به شیوه بازنمایی پارامترهای زمان و دامنه در یک شکل موج است. شکل ۱-۳ تفاوت بین سیگنال‌های گسسته و پیوسته را نشان می‌دهد (برای آنکه هر دو موج به آسانی قابل مشاهده باشند، با اندکی اختلاف فاز نشان داده شده‌اند). همان‌طور که در شکل می‌بینید، یک سیگنال پیوسته مانند خط پیوسته‌ای است که در همه نقاط زمان دارای مقدار معینی دامنه است. اما یک سیگنال گسسته در واقع، توالی مقادیر گسسته دامنه است و به همین دلیل، به صورت نقاطی جدا از هم (گسسته)، نه یک خط پیوسته، نشان داده شده است. زمان و دامنه در یک سیگنال پیوسته، به صورت اعدادی بازنمایی می‌شوند که به لحاظ نظری، بی‌نهایت رقم بعد از نقطه اعشار دارند. به عنوان مثال، ممکن است مقدار دامنه یک سیگنال را $\dots 3/211178$ ولت، در زمان $\dots 1/00333483$ ثانیه اندازه‌گیری کنیم، اما در سیگنال‌های گسسته، تعداد رقم‌های اعشار بعد از نقطه اعشار (برای هر دو مقدار زمان و دامنه) همواره محدود است.



شکل ۱-۳. امواج سینوسی پیوسته و گسسته.

بسیاری از ابزارهایی که برای ضبط و تجزیه گفتار به کار می‌بریم، می‌توانند تغییرات پیوسته فشار هوا را به سیگنال پیوسته الکتریکی تبدیل کنند، اما کامپیوترها این قابلیت را ندارند. کامپیوترها نمی‌توانند اعداد را با شمار نامحدودی رقم‌های اعشاری ذخیره کنند. بنابراین، همه سیگنال‌های آکوستیکی در کامپیوترها به صورت گسسته ذخیره می‌شوند هرچند که نوسانات فشاری که تولید صدا می‌کنند، همواره پیوسته هستند. وسایلی مانند ضبط صوت که می‌توانند یک سیگنال پیوسته را ذخیره کنند، ابزار آنالوگ^۱ نامیده می‌شوند، زیرا می‌توانند سیگنال آکوستیکی را به صورت یک سیگنال آنالوگ الکتریکی ذخیره کنند. اما وسایلی که سیگنال پیوسته گفتاری را به یک سیگنال گسسته تبدیل می‌کنند، ابزار دیجیتال^۲ یا رقمی نامیده می‌شوند، زیرا در این ابزارها، سیگنال به صورت دیجیت^۳ یا رقم ذخیره می‌شود. از این رو، واژه‌های «آنالوگ» و «دیجیتال» مترادف واژه‌های پیوسته و گسسته است. قابل ذکر است که ابزارهای آنالوگ ضرورتاً برتری در ذخیره و تجزیه صدا ندارند. اگر هنگام تبدیل یک سیگنال پیوسته به گسسته، رقم‌های اعشاری را به اندازه لازم حفظ کنید، همه اطلاعات شنیداری مهم ذخیره می‌شود (همان‌طور که به عنوان مثال، در دیسک‌های فشرده^۴ صورت می‌گیرد).

۳-۲. تبدیل آنالوگ به دیجیتال

برای «ضبط» گفتار به وسیله یک کامپیوتر، لازم است که موج صوتی پیوسته به یک سیگنال گسسته تبدیل شود تا قابل ذخیره (و کاربرد) به وسیله کامپیوتر باشد. تبدیل موج صوتی پیوسته به گسسته طی یک فرایند دو مرحله‌ای به نام تبدیل آنالوگ به دیجیتال صورت می‌گیرد. در مرحله اول، تعداد رقم‌های اعشاری پس از نقطه اعشار بر روی محور زمان محدود می‌شود. این عمل سبب می‌شود که سیگنال پیوسته به جای آنکه از بی‌نهایت نقطه در زمان تشکیل شده باشد، به توالی از «نمونه‌ها» یا

1. analog devices

2. digital devices

3. digit

4. compact disc

نقاط گسسته در زمان تقطیع شود. در مرحله بعد، تعداد رقم‌های اعشار پس از نقطه اعشار بر روی محور دامنه موج محدود می‌شود. بنابراین، سیگنال به جای آنکه از بی‌نهایت سطوح دامنه تشکیل شده باشد، کمی^۱ می‌شود. برای آنکه بازنمایی دیجیتالی دقیقی از سیگنال آکوستیکی به دست آید، لازم است به برخی مسائل محتمل که ممکن است در هر یک از مراحل نمونه‌برداری و کمی‌سازی در تبدیل آنالوگ به دیجیتال روی دهد، توجه نمود.

مثالی برای تبدیل آنالوگ به دیجیتال

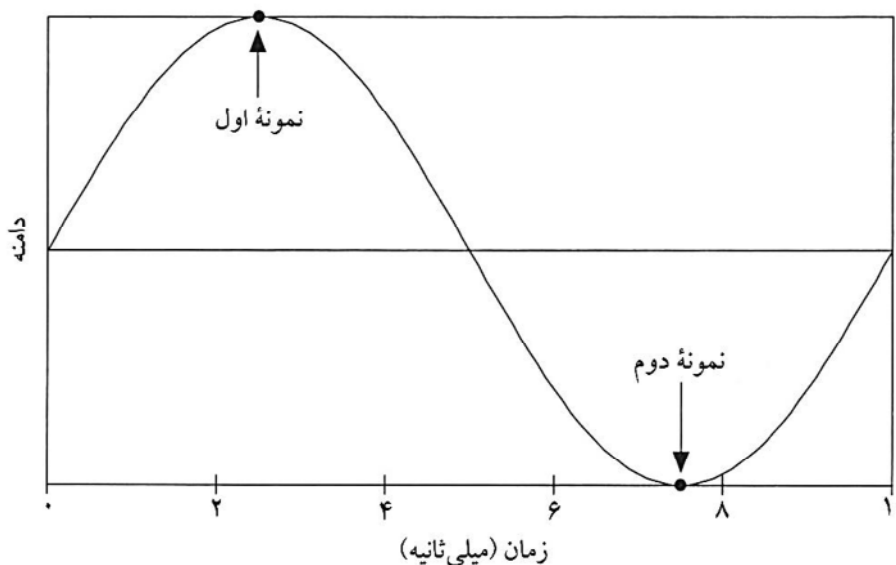
تبدیل آنالوگ به دیجیتال شبیه تبدیل نمودار یک موج سینوسی ساده به جدولی از اعداد با استفاده از یک خط‌کش است. ابتدا فواصل یکسانی را روی محور زمان علامت‌گذاری کنید. سپس، در هر یک از این نقاط، مقادیر دامنه موج را نیز اندازه‌گیری کنید. علامت‌گذاری فواصل مساوی بر روی محور زمان، نمونه‌برداری نامیده می‌شود که در بخش ۳-۲-۱ مورد بحث قرار می‌گیرد و اندازه‌گیری مقادیر دامنه در هر یک از نقاط زمانی، کمی‌سازی نامیده می‌شود که در بخش ۳-۲-۲ مطرح می‌شود.

۳-۲-۱. نمونه‌برداری

مهم‌ترین نکته‌ای را که در مرحله نمونه‌برداری تبدیل آنالوگ به دیجیتال باید مورد توجه قرار داد، نرخ نمونه‌برداری^۲ است. نرخ نمونه‌برداری عبارت است از تعداد دفعاتی در ثانیه که موج پیوسته را برای تولید موج گسسته اندازه‌گیری می‌کنیم. بسیار مهم است که نمونه‌برداری از سیگنال به اندازه کافی صورت گیرد تا همه اطلاعات طیفی که قصد مطالعه آن را داریم، در سیگنال دیجیتال ذخیره شود. فرض کنید یک موج سینوسی داریم که ۱۰۰ بار در ثانیه تکرار می‌شود (۱۰۰ هرتز). چند نمونه در ثانیه باید برداریم (با فرض اینکه فاصله زمانی بین نمونه‌ها ثابت است) تا در نهایت،

1. quantized 2. sampling rate

این واقعیت که یک موج متناوب ۱۰۰ هرتزی در سیگنال پیوسته وجود دارد، ذخیره شود؟ همان‌طور که شکل ۲-۳ نشان می‌دهد، حداقل ۲ نمونه برای هر سیکل موج سینوسی لازم است تا الگوی تناوبی^۱ موج ذخیره شود. هرچند که برای دانستن دامنه و فاز موج، نمونه‌های بیشتری لازم است، اما دو نمونه در هر سیکل کافی است تا به ما بگوید که یک مؤلفه با آن فرکانس در سیگنال وجود دارد. پس برای ذخیره یک مؤلفه متناوب ۱۰۰ هرتزی در یک موج پیوسته، بازنمایی گسسته آن باید با حداقل دو نمونه برای هر سیکل موج پیوسته صورت گیرد. به عبارت دیگر، اگر بخواهیم مؤلفه‌های سیگنال را تا ۱۰۰ هرتز تجزیه کنیم، نرخ نمونه‌برداری باید حداقل ۲۰۰ هرتز باشد.



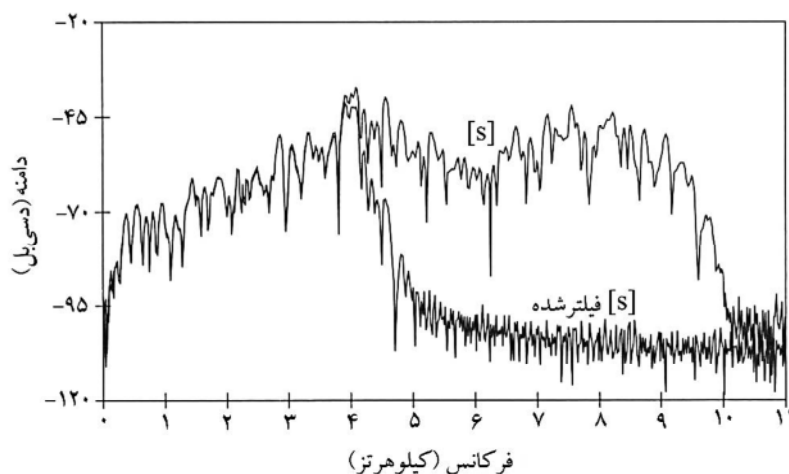
شکل ۲-۳. بازنمایی نشان می‌دهد که چرا دو نقطه برای ذخیره الگوی تناوبی موج سینوسی لازم است. اصطلاحی که اغلب در این زمینه مطرح می‌شود، فرکانس نایکویست^۲ است. فرکانس نایکویست بالاترین فرکانسی است که با یک نرخ نمونه‌برداری معین می‌توان ذخیره کرد. در مثال بالا با نرخ نمونه‌برداری ۲۰۰ هرتز، فرکانس نایکویست ۱۰۰ هرتز است. فرکانس نایکویست، نصف نرخ نمونه‌برداری است.

1. periodicity

2. Nyquist frequency

پردازش دیجیتالی سیگنال و سی‌دی‌ها^۱

اگر اخیراً برای خرید یک دستگاه استریو به فروشگاه‌های لوازم صوتی سر زده باشید، اغلب با این ادعای فروشندگان مواجه شده‌اید که پاسخ فرکانسی بلندگوها بین ۲۰ هرتز تا ۲۰۰۰۰ هرتز است. آنها معمولاً این محدوده فرکانسی را به کار می‌برند، زیرا این محدوده پایین‌ترین و بالاترین حد فرکانسی است که انسان می‌تواند بشنود (در واقع همان‌طور که انسان مسن‌تر می‌شود، حساسیت گوش او معمولاً به فرکانس‌های بالا کاهش می‌یابد و فرکانس ۲۰۰۰۰ هرتز بالاترین حد شنوایی گوش‌های جوان و سالم است). از آنجا که نرخ نمونه‌برداری در تبدیل آنالوگ به دیجیتال باید دو برابر فرکانس هر مؤلفه طیفی باشد که قصد ذخیره آن را داریم، پس اگر هدف ما ذخیره همه مؤلفه‌های قابل شنیدن یک موج (از ۲۰ هرتز تا ۲۰۰۰۰ هرتز) است، باید نرخ نمونه‌برداری در حدود ۴۰۰۰۰ هرتز باشد. به همین دلیل نرخ نمونه‌برداری در ابزارهای صوتی دیجیتالی مثل ابزارهای سی‌دی خوان^۲، ۴۴۰۰۰ هرتز می‌باشد.



شکل ۳-۳. طیف‌های قدرت [S] که نشان می‌دهد این آوا، مؤلفه‌های طیفی با دامنه نسبتاً زیاد در فرکانس‌های بالای ۴ کیلوهرتز دارد. هر دو طیف تقریباً تا ۴ کیلوهرتز یکسان هستند. شکل موج طیفی که عنوان «[S] فیلتر شده» دارد، با یک صافی پایین‌گذر با فرکانس قطع ۴ کیلوهرتز فیلتر شده است. شکل موج طیفی که عنوان «[S]» دارد، نیز فیلتر شده اما با یک صافی پایین‌گذر که فرکانس قطع آن ۹ کیلوهرتز است.

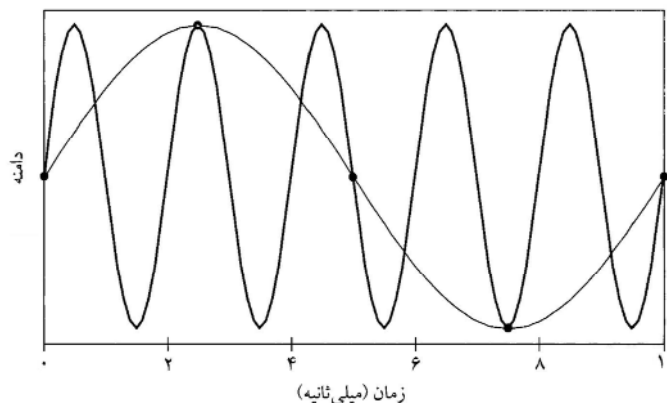
1. CD

2. CD player

بین نرخ نمونه‌برداری و میزان فضایی که یک موج صوتی در حافظه کامپیوتر اشغال می‌کند، رابطه بده‌بستان وجود دارد. اگر یک موج را با نرخ ۴۰۰۰۰ هرتز (۴۰ کیلوهرتز) نمونه‌برداری کنیم، باید ۴۰۰۰۰ نمونه برای هر ثانیه از گفتار ذخیره کنیم. اگر قصد دارید گفتار را برای ساعت‌های طولانی ذخیره کنید، بهتر است یک نرخ نمونه‌برداری پایین‌تر به کار ببرید. تلفن محدودیت باند فرکانسی دارد، بدین صورت که فقط مؤلفه‌های بالای ۳۰۰ هرتز و زیر ۴۰۰۰ هرتز را انتقال می‌دهد. این پهنای باند، به‌جامانده از روزگار استفاده از میکروفون‌های کربنی است که چنین پاسخ فرکانسی داشتند و به این معناست که می‌توان گفتار را با نرخ فقط ۸۰۰۰ نمونه در ثانیه (۸ کیلوهرتز) نمونه‌برداری کرد و گفتار ضبط‌شده مانند گفتار روی خط تلفن قابل فهم باشد. اما همان‌طور که شکل ۳-۳ نیز نشان می‌دهد، برخی آواهای گفتاری (مثل [s]) معمولاً میزان قابل توجهی انرژی در فرکانس‌های بالای ۴۰۰۰ هرتز دارند. پژوهش‌هایی که در زمینه ویژگی‌های آکوستیکی گفتار صورت گرفته، نشان می‌دهد که هیچ آوای گفتاری نیست که در فرکانس‌های بالای ۱۰ کیلوهرتز اطلاعات آکوستیکی معناداری داشته باشد. بنابراین، یک نرخ نمونه‌برداری در حدود ۲۰ کیلوهرتز برای ضبط گفتار مناسب است. بدیهی است که انتخاب سازندگان استریو برای طراحی سیستم صوتی برای گوش‌های «جوان و سالم» انتخابی بسیار کشنده برای جامعه گفتاری محسوب می‌شود. یکی از دلایل آن، این است که جامعه گفتاری دربرگیرنده طیف گسترده‌ای از گویشوران (پیر و جوان) است و در محیط‌های گفتاری بسیار متنوعی روی می‌دهد (که برخی تا حدی نوفه‌ای هستند). همان‌طور که در فصل بعد خواهیم دید، به نظر نمی‌رسد که مؤلفه‌های فرکانسی بالای ۱۰ کیلوهرتز برای جامعه گفتاری مفید باشند، حتی اگر شنوندگان از شنوایی کامل و سالمی برخوردار باشند، زیرا حساسیت شنوایی انسان به اطلاعات فرکانسی بالای ۱۰ کیلوهرتز نسبتاً محدود است. از این رو، پیشنهاد می‌کنم به جای نمونه‌برداری گفتار در ۴۴ کیلوهرتز، فرکانس ۲۲ کیلوهرتز را به عنوان نرخ نمونه‌برداری پیش‌فرض خود به کار ببرید.

شکل ۳-۴ مسئله دیگری را نشان می‌دهد که هنگام تبدیل یک سیگنال پیوسته به یک سیگنال گسسته روی می‌دهد. در این مثال، یک موج سینوسی پیوسته داریم که

سرعت نوسان آن به قدری زیاد است که نمی‌توان آن را به‌طور کامل با نمونه‌های گسسته بازنمایی کرد. فرکانس سیگنال پیوسته بزرگ‌تر از نصف نرخ نمونه‌برداری (فرکانس نایکویست) است. در نتیجه، موج گسسته حاصل، فرکانس کاملاً متفاوتی نسبت به سیگنال پیوسته اولیه دارد. مسئله این است که سیگنال پیوسته بیش از یک سیکل را بین نمونه‌های متوالی موج گسسته کامل می‌کند. در نتیجه، به نظر می‌رسد که گویا نمونه‌برداری از موجی با فرکانس خیلی پایین‌تر صورت گرفته است. بازنمایی نادرست یک سیگنال پیوسته در تبدیل به موج گسسته، **بدنمایی**^۱ نامیده می‌شود و هنگامی روی می‌دهد که سیگنال پیوسته مؤلفه‌هایی دارد که فرکانس آنها بالاتر از نصف نرخ نمونه‌برداری است. برای اهتزاز از بدنمایی، یا باید نرخ نمونه‌برداری را برای ذخیره همه مؤلفه‌های فرکانسی سیگنال پیوسته افزایش دهیم، یا همه فرکانس‌های بالای نصف نرخ نمونه‌برداری را فیلتر کنیم. از آنجا که در سیگنال‌های طبیعی (برخلاف امواج سینوسی خالص)^۲ همیشه مؤلفه‌هایی با فرکانس بسیار بالا وجود دارد، همواره ضروری است که فرکانس‌های بالای نصف نرخ نمونه‌برداری فیلتر شوند. بیشتر کارت‌های صوتی کامپیوترها این کار را به‌طور خودکار انجام می‌دهند.



شکل ۳-۴. بازنمایی بدنمایی در تبدیل آنالوگ به دیجیتال. سرعت نوسان سیگنال آنالوگ (خط پررنگ) به قدری زیاد است که نمی‌توان آن را دقیقاً با نمونه‌های سیگنال دیجیتال (نقاط توپر) بازنمایی کرد. از این رو، سیگنال دیجیتال نه تنها بازنمایی دقیقی از سیگنال ورودی که فرکانس بالا دارد، نیست، بلکه حاوی مؤلفه‌ای با فرکانس پایین (خط کم‌رنگ) است.

1. aliasing 2. pure sine wave

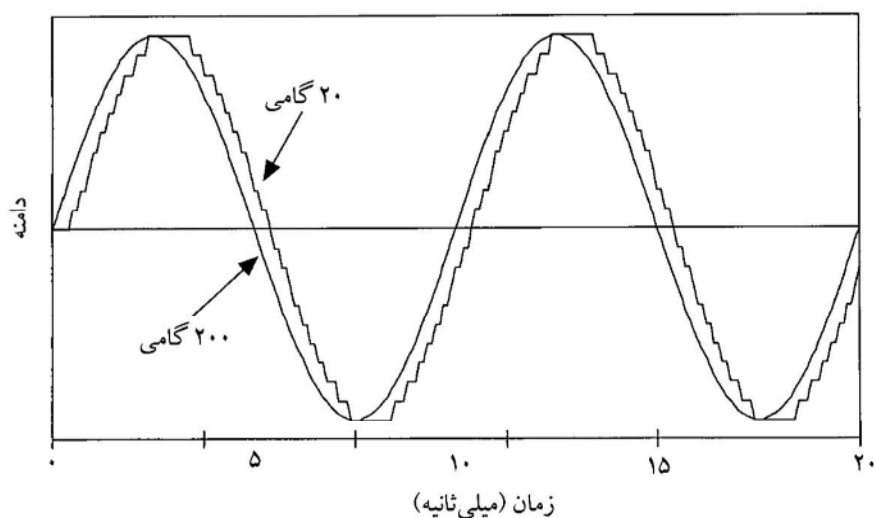
همان‌طور که گفته شد، می‌توان همه اطلاعات شنیداری مهم در سیگنال گفتار را با انتخاب یک نرخ نمونه‌برداری در حدود ۲۰ کیلوهرتز ذخیره کرد، زیرا بدین‌طریق همه مؤلفه‌های سیگنال از صفر تا ۱۰ کیلوهرتز ذخیره می‌شوند. همان‌طور که شکل ۳-۴ نیز نشان می‌دهد، هر مؤلفه‌ای با فرکانس بالاتر از ۱۰ کیلوهرتز نه‌تنها در موج دیجیتال ذخیره نمی‌شود، بلکه با ایجاد مؤلفه‌های بدنما در سیگنال گسسته، سبب خراب‌شدن آن نیز می‌شود. از این رو، استفاده از یک صافی پایین‌گذر برای حذف همه مؤلفه‌های فرکانسی سیگنال پیوسته که تولید مؤلفه‌های بدنما می‌کنند، ضروری است. این صافی، صافی پادبدنمایی^۱ نامیده می‌شود. اگر نرخ نمونه‌برداری ۲۰ کیلوهرتز باشد، یک صافی پادبدنمایی برای حذف همه مؤلفه‌های فرکانسی بالای ۱۰ کیلوهرتز و عبور همه مؤلفه‌های فرکانسی زیر ۱۰ کیلوهرتز مورد نیاز است. از آنجا که شیب صافی‌های واقعی کاملاً تند نیست، در واقع باید از یک صافی پایین‌گذر با فرکانس قطعی در حدود ۷ یا ۷/۵ کیلوهرتز برای نرخ نمونه‌برداری ۲۰ کیلوهرتز استفاده کنیم. در این صورت می‌توانیم مطمئن باشیم که مؤلفه‌های فرکانسی که تولید بدنمایی می‌کنند، به‌وسیله صافی حذف می‌شوند. ضبط صوت‌های دیجیتال معمولاً سیگنال را با نرخ ۴۴ کیلوهرتز نمونه‌برداری و در ۲۰ کیلوهرتز فیلتر پایین‌گذر می‌کنند. وقتی که شما صوت را به‌وسیله میکروفون یا پورت صوتی کامپیوتر خود ضبط می‌کنید، می‌توانید یک نرخ نمونه‌برداری انتخاب کنید (من معمولاً ۲۲ کیلوهرتز را انتخاب می‌کنم) و کارت صوتی کامپیوتر به‌طور خودکار سیگنال را در فرکانسی دقیقاً زیر فرکانس نایکویست فیلتر پایین‌گذر می‌کند.

۳-۲-۲. کمی‌سازی

پس از آنکه یک سیگنال پیوسته را بر روی محور زمان به نقاط گسسته تقطیع کردیم، باید دامنه سیگنال را نیز در هر یک از این نقاط اندازه‌گیری کنیم. این مرحله از تبدیل آنالوگ به دیجیتال، یعنی کمی‌سازی، دقیقاً شبیه اندازه‌گیری دامنه موج با یک خط‌کش است.

1. anti-aliasing filter

مهم‌ترین موضوعی که در کمی‌سازی باید مدنظر داشت، میزان دقت اندازه‌گیری دامنه است. میزان دقت اندازه‌گیری وقتی مطرح می‌شود که دامنه را با استفاده از یک خط‌کش اندازه‌گیری می‌کنید. آیا مقادیر اندازه‌گیری را به نزدیک‌ترین $\frac{1}{4}$ اینچ گرد می‌کنید یا به نزدیک‌ترین $\frac{1}{8}$ اینچ؟ شکل ۳-۵ تأثیر انتخاب دو مقیاس اندازه‌گیری متفاوت را نشان می‌دهد. یک بار دامنه با ۲۰ مقدار و بار دیگر با ۲۰۰ مقدار ممکن بازنمایی شده است. (برای رسم این نمودار، مقادیر حاصل از تابع سینوسی را که به‌طور پیوسته از ۱ تا ۱- متغیرند، یا تا یک رقم اعشار (۰، ۰/۲، ۰/۱، ۰) برای به‌دست‌آوردن کمی‌سازی ۲۰ گامی، یا تا دو رقم اعشار (۰، ۰/۰۲، ۰/۰۱، ۰) برای به‌دست‌آوردن کمی‌سازی ۲۰۰ گامی گرد کردم). همان‌طور که شکل ۳-۵ نیز نشان می‌دهد، اگر دامنه با دقت لازم اندازه‌گیری نشود، یک موج پیوسته هموار، ممکن است در نهایت دندان‌دندانه به نظر برسد. در این بخش به‌طور خلاصه خواهیم گفت که منظور از «دقت لازم» چیست.



شکل ۳-۵. بازنمایی دو الگوی متفاوت کمی‌سازی. در موج سینوسی که با عنوان «گامی ۲۰» مشخص شده، دامنه با ۲۰ مقدار ممکن بازنمایی شده است. دامنه همین موج سینوسی در نمونه‌ای که با عنوان «گامی ۲۰۰» مشخص شده، با ۲۰۰ مقدار ممکن بازنمایی شده است. برای آنکه مقایسه دو موج به راحتی صورت گیرد، با اندکی اختلاف فاز نشان داده شده‌اند.

برای تعریف «دقت لازم» به یک واحد اندازه‌گیری نیاز داریم. مثلاً برای کمی‌سازی یک موج چاپ‌شده، یک واحد اندازه‌گیری مثل اینچ یا میلی‌متر (یا پیکسل^۱) برای تعیین میزان «دقت لازم» انتخاب می‌کنیم. در تبدیل رایانه‌ای موج آنالوگ به دیجیتال نیز به‌طور مشابه روش‌های متعددی برای اندازه‌گیری دقت کمی‌سازی وجود دارد. مثلاً می‌توان «دقت لازم» را برحسب فشار صوت یا برحسب ولتاژ تولیدشده به‌وسیله یک میکروفون تعیین کرد. اما هر دوی این واحدهای محتمل اندازه‌گیری به‌قدری با فرایند کمی‌سازی تفاوت دارند که کاربرد عملی چندانی ندارند. در عمل، دقت کمی‌سازی بر اساس تعداد سطوح مستقل دامنه که بتوان آنها را به‌صورت دیجیتالی در یک کامپیوتر بازنمایی کرد، تعیین می‌شود (در این مورد خط‌کش مثال خوبی است، زیرا «تعداد سطوح مستقل دامنه که بتوان آنها را بازنمایی کرد»، شبیه به اندازه‌گیری تا $\frac{1}{4}$ اینچ یا تا $\frac{1}{8}$ اینچ است. دقت اندازه‌گیری به این بستگی دارد که به چه میزان تقسیم‌بندی صورت می‌گیرد). از آنجا که کامپیوترها با اعداد صحیح (۱، ۲، ۳ و ...) خیلی بهتر می‌توانند کار کنند تا با اعداد اعشاری (۰/۰۱، ۰/۰۲، ۰/۰۳ و ...) و نیز از سویی فرایند کمی‌سازی مستلزم تقلیل یک مقیاس پیوسته دامنه به یک مقیاس گسسته با استفاده از شمار محدودی مقادیر ممکن دامنه است، بنابراین، کامپیوترها موج آکوستیکی را به‌صورت توالی از اعداد صحیح ذخیره می‌کنند. مثلاً در کمی‌سازی ۲۰ گامی که در شکل ۳-۵ نشان داده شد، کامپیوتر به‌جای ذخیره اعداد اعشاری مانند ۰/۸، ۰/۹ و غیره، اعداد صحیح از ۱ تا ۲۰ (یا از ۱۰- تا ۱۰+) را ذخیره می‌کند. اندازه احتمالی یک عدد صحیح نیز بر اساس تعداد رقم‌های دوتایی (بیت‌ها^۲) که برای بازنمایی آن در حافظه کامپیوتر به کار می‌روند، تعیین می‌شود (به مستطیل «رقم‌های دیجیتالی» رجوع کنید). هر قدر تعداد بیت‌های به‌کاررفته برای بازنمایی هر نمونه در یک موج دیجیتالی بیشتر باشد، وضوح دامنه^۳ نیز بیشتر خواهد بود. رایج‌ترین تعداد بیت‌های به‌کاررفته برای کدبندی نمونه‌های

گفتاری ۸، ۱۲ و ۱۶ تایی است.

1. pixel

2. bit

3. amplitude resolution

رقم‌های دیجیتالی

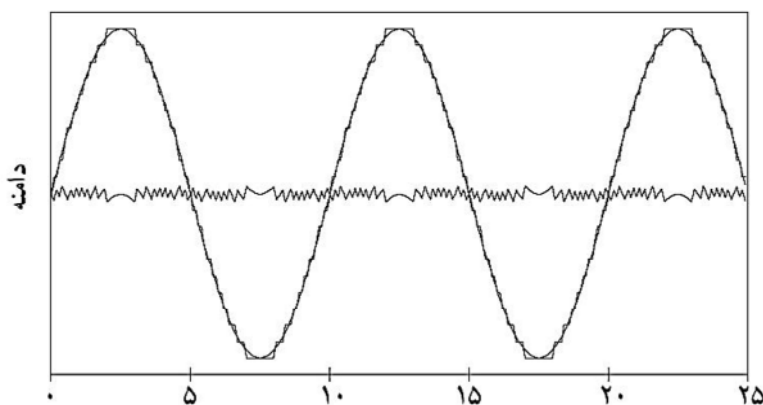
کامپیوترها برای ذخیره اطلاعات از توالی صفر و یک استفاده می‌کنند. تعداد این رقم‌های دوتایی (بیت‌ها) که برای کدبندی یک عدد صحیح به کار می‌رود، حداکثر مقدار آن عدد را تعیین می‌کند و بنابراین، در یک شکل موج دیجیتال، تعداد بیت‌ها تعیین‌کننده درجه وضوح دامنه انرژی است. مثلاً یک رقم دوبیتی می‌تواند یکی از چهار ترکیب مختلف ۱ و ۰ به صورت (۱۱، ۱۰، ۰۱، ۰۰) باشد. اعداد صحیح حاصل از این ترکیب‌ها عبارتند از: $۱۱ = ۳$ ، $۱۰ = ۲$ ، $۰۱ = ۱$ ، $۰۰ = ۰$. پس با رقم‌های دوبیتی، کامپیوتر می‌تواند اعداد صحیح از صفر تا ۳ را بشمارد. یک رقم سه‌بیتی می‌تواند یکی از هشت ترکیب مختلف صفرها و یک‌ها باشد که اعداد صحیح از صفر تا ۷ را نشان می‌دهد و یک رقم چهاربیتی می‌تواند یکی از $۲^4 = ۱۶$ ترکیب مختلف صفرها و یک‌ها باشد که نشان‌دهنده اعداد صحیح از صفر تا ۱۵ است. بدین ترتیب، تعداد بیت‌های به کاررفته برای ذخیره اعداد صحیح در یک کامپیوتر شبیه به مقیاس‌های روی یک خط‌کش است. از این رو، دقت کمی‌سازی در تبدیل آنالوگ به دیجیتال معمولاً برحسب تعداد بیت‌های به کاررفته برای ذخیره نمونه‌های شکل موج تعیین می‌شود.

فرایند تبدیل موج آنالوگ به دیجیتال سبب ایجاد مقداری خطا در سیگنال دیجیتال می‌شود، زیرا در واقع، یک سیگنال پیوسته در حال تبدیل شدن به یک سیگنال گسسته است. شکل ۳-۶ این واقعیت را نشان می‌دهد. در این شکل یک موج پیوسته، بازنمایی گسسته آن (با دامنه ۲۰ گامی) و نیز شکل موجی که اختلاف بین شکل موج‌های پیوسته و گسسته را نشان می‌دهد، مشاهده می‌شود. شکل موج سوم، نوفه کمی‌سازی^۱ در شکل موج گسسته را نشان می‌دهد. همه امواج دیجیتال با مقداری نوفه کمی‌سازی همراه هستند، زیرا در این امواج، سیگنال به جای آنکه به طور

1. quantization noise

در کمی‌سازی، مقادیر دامنه موج پیوسته که بین دو گام کمی‌سازی قرار دارند، به نزدیک‌ترین گام گرد می‌شوند. این امر سبب به وجود آمدن مقداری خطا در سیگنال گسسته حاصل می‌شود. به عبارت دیگر، مقدار گسسته برابر است با مقدار پیوسته + مقداری خطا. این خطا به صورت نوفه، همراه با سیگنال گسسته تولید می‌شود که به آن نوفه کمی‌سازی می‌گویند. (مترجم)

پیوسته از همه نقاط ممکن زمان عبور کند، از یک گام کمی سازی به گام بعدی می پرد. شکل موج نوفه کمی سازی شبه تصادفی است (نوعی شبیه به نوفه سفید) و در محدوده ای بین $\frac{1}{4}$ تا $\frac{1}{2}$ گام های دامنه نوسان می کند. به عبارت دیگر، دامنه در شکل موج پیوسته همواره بین $\frac{1}{4}$ تا $\frac{1}{2}$ گام دامنه ای است که در شکل موج گسسته ضبط می شود (مشروط بر آنکه یک صافی پادبدنمایی مناسب به کار رود). گاهی اوقات ممکن است نوفه کمی سازی را هنگام پخش یک صوت دیجیتال بشنوید (به ویژه با کمی سازی ۸ بیتی). بهترین راه برای جلوگیری از تولید آن، افزایش گام های دامنه است (نمودار ۲۰ گامی را با نمودار ۲۰۰ گامی مقایسه کنید). دامنه نوفه کمی سازی را می توان در مقایسه با ماکزیمم دامنه سیگنال بیان کرد. اگر گام های دامنه ۲۵۶ تا باشد (نمونه های ۸ بیتی)، نسبت ماکزیمم دامنه سیگنال به دامنه نوفه کمی سازی ۲۵۶:۱ است، اما در نمونه های ۱۲ بیتی، این نسبت ۴۰۹۶:۱ و در نمونه های ۱۶ بیتی ۶۵۵۳۶:۱ است. بی تردید هنگامی که بیت های بیشتری برای ذخیره سیگنال به کار می رود، بلندی نسبی نوفه کمی سازی (نسبت سیگنال به نوفه) پایین تر است. به همین دلیل، در ابزارهای سی دی خوان، کمی سازی ۱۶ بیتی و در برخی تبدیل کننده های آنالوگ به دیجیتال، نمونه برداری ۲۴ بیتی به کار می رود.



شکل ۳-۶. بازنمایی نوفه کمی سازی. این شکل بازنمایی پیوسته و گسسته یک موج سینوسی را نشان می دهد (دو موج، هم فاز هستند و بنابراین، روی هم قرار گرفته اند). سیگنال تصادفی که حول دامنه صفر نوسان می کند، اختلاف بین دو بازنمایی پیوسته و گسسته را نشان می دهد: نوفه کمی سازی.

در گذشته، پژوهشگران کمی سازی ۸ بیتی را برای ذخیره دادگان صوتی حجیم (روی فلاپی دیسک‌ها) به کار می‌بردند. اما امروزه معمولاً دادگان حجیم با فرمت ۱۶ بیتی بر روی سی‌دی‌رام‌ها^۱ یا اینترنت ذخیره می‌شود. فشردن سازی داده‌ها نیز برای کاهش حجم داده‌های صوتی بسیار رایج است. آشناترین نوع فشردن سازی داده‌های صوتی، الگوریتم فشردن سازی ام‌پی‌تری^۲ است که برای پخش برخت موسیقی در اینترنت به کار می‌رود. دانستن تفاوت بین فشردن سازی «بافت» و «بی‌افت» بسیار مهم است. برنامه‌های فشردن سازی همراه با افت، اطلاعات را حذف می‌کنند، به نحوی که نمی‌توان به‌طور دقیق موج دیجیتال اولیه را از داده‌های فشردن شده دوباره بازسازی کرد. به عنوان مثال، ام‌پی‌تری یک برنامه فشردن سازی همراه با افت است و به همین دلیل، صوت ضبط شده با ام‌پی‌تری به خوبی صوت اولیه نیست (هرچند که ام‌پی‌تری استاندارد اصلی پخش موسیقی در اینترنت است). جدول ۳-۱ برخی از برنامه‌های متداول فشردن سازی بافت و بی‌افت را نشان می‌دهد.

جدول ۳-۱. استانداردهای فشردن سازی بافت و بی‌افت برای فایل‌های تصویری، داده‌های صوتی و فایل‌های کامپیوتری

بی‌افت	بافت	
GIF, PNG, TIFF	JPEG	فایل‌های تصویری
WAV, AIFF	MP3	داده‌های صوتی
ZIP, TAR		سایر فایل‌های کامپیوتری

از آنجا که دامنه نوفه کمی سازی ثابت است، پس دامنه سیگنال، تعیین کننده نسبت سیگنال به نوفه است. از این رو، هنگام دیجیتالی کردن گفتار (تبدیل آنالوگ به دیجیتال) بسیار مهم است که دامنه سیگنال را تا آنجا که ممکن است بالا نگه داریم، بی‌آنکه از محدوده دامنه‌ای که کامپیوتر می‌تواند بپذیرد، خارج شویم. مثلاً اگر مبدل آنالوگ به دیجیتال کامپیوتر به محدوده دامنه‌ای بین ۱۰- تا ۱۰ ولت حساس است و محدوده دامنه سیگنال ورودی به کامپیوتر بین ۵- تا ۵ ولت باشد، در این

1. CD-ROM 2. MP3

صورت، سیگنال دیجیتال حاصل نوفه کمی سازی نسبتاً بالاتری خواهد داشت در مقایسه با وقتی که شما دامنه سیگنال را به طور صحیحی تقویت^۱ می کنید تا محدوده دامنه ای که کامپیوتر می تواند بپذیرد، را پر کند. در این مثال، با یک مبدل آنالوگ به دیجیتال ۸ بیتی و استفاده از نصف محدوده ولتاژ ورودی، فقط از ۷ تا ۸ بیت ممکن برای ذخیره هر نمونه استفاده می کنیم و بنابراین، نسبت سیگنال به نوفه ۱:۱۲۸ است و نه ۱:۲۵۶. نسبت سیگنال به نوفه در مبدل های آنالوگ به دیجیتال ۱۶ بیتی معمولاً مسئله مهمی نیست، زیرا سطح نوفه کمی سازی با نمونه برداری ۱۶ بیتی نسبتاً پایین است، اما هرگز نباید هنگام دیجیتالی کردن گفتار از نسبت سیگنال به نوفه غفلت ورزید. با این حال، به یک اخطار نیز توجه کنید: اگر ولتاژ ورودی بیش از محدوده ولتاژی باشد که می تواند در شکل موج دیجیتال بازنمایی شود، نتایج نامطلوبی خواهد داشت که به آن **برش سیگنال**^۲ گفته می شود. برش سیگنال سبب تولید امواج گذرا در شکل موج دیجیتال می شود و در نتیجه، سبب خراب شدن هر تجزیه طیفی می گردد که قصد دارید روی سیگنال انجام دهید (طیف گذرای شکل ۱۱-۱ را ملاحظه کنید).

۳-۳. روش های تجزیه سیگنال

در این بخش شش روش تجزیه در پردازش دیجیتال سیگنال که معمولاً در آواشناسی آکوستیک به کار می رود، معرفی می شوند. هدف من ارائه برخی اطلاعات پایه درباره این روش هاست، به گونه ای که پس از آن خواننده بتواند پارامترهای تجزیه را در نرم افزار تجزیه گفتار خود تنظیم کند. از این رو، مطالب این بخش مقدمه عملی برای این روش هاست، نه توضیح مبسوط ریاضی آنها که در کتاب هایی مثل اشگنسی^۳ (۱۹۸۷) یا سایر کتاب های مهندسی ملاحظه می شود.

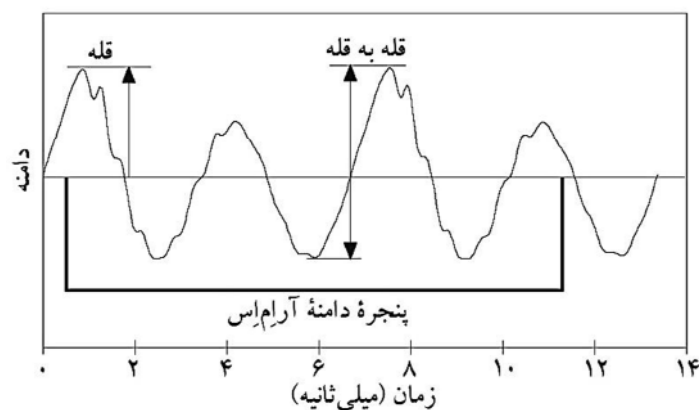
1. amplify

2. signal clipping

3. O'Shaughnessy

۳-۱-۳. دامنه آرام‌اس^۱

همان‌طور که در شکل ۷-۳ می‌بینید، دامنه یک شکل‌موج را می‌توان به روش‌های مختلفی اندازه‌گیری کرد. در امواج متناوب ساده (سینوسی)، سه روش اندازه‌گیری دامنه، یعنی دامنه قله، دامنه قله‌به‌قله و دامنه آرام‌اس (دامنه ریشه میانگین مربعات) با روش‌های ریاضی از یکدیگر مشتق می‌شوند. مثلاً دامنه قله‌به‌قله دوبرابر دامنه قله است و دامنه آرام‌اس در یک سیکل موج سینوسی برابر با دامنه قله ضربدر عدد 0.707 است. اما در امواج مرکب روش‌های اندازه‌گیری مختلف دامنه از یکدیگر قابل پیش‌بینی نیست. روش‌های اندازه‌گیری دامنه برحسب قله [دامنه قله و دامنه قله‌به‌قله] با دامنه آرام‌اس تفاوت دارد، زیرا این روش‌ها میزان دامنه آکوستیکی را نشان می‌دهند، حال آنکه دامنه آرام‌اس، نشان‌دهنده میزان شدت آکوستیکی است. از آنجا که بلندی درکی صدا بیشتر به شدت آکوستیکی ارتباط دارد تا به دامنه آکوستیکی، بیشتر نرم‌افزارهای تجزیه صوتی به‌طور خودکار دامنه آرام‌اس را محاسبه و گزارش می‌کنند.



شکل ۷-۳. نمایش سه روش اندازه‌گیری دامنه در دو دوره تناوب یک‌واکه: دامنه قله، دامنه قله‌به‌قله و پنجره‌ای که دامنه آرام‌اس را می‌توان در آن به دست آورد.

1. root mean square (RMS)

برای محاسبه دامنه آرام‌اس، ابتدا مجذور هر نمونه در یک «پنجره» شکل موج به دست می‌آید (به مطلب «پنجره‌گذاری»^۱ رجوع کنید). سپس میانگین مربعات محاسبه می‌شود و در نهایت، ریشه میانگین مربعات به دست می‌آید. این روش به‌طور خلاصه، ریشه میانگین مربعات نامیده می‌شود. به‌عنوان مثال اگر پنجره‌ای با ۹ نمونه به‌صورت $(0, -2, 0, -5, -3, 0, 2, 5, 0)$ داشته باشیم، مجذور این مقادیر به‌صورت $(0, 4, 0, 25, 9, 0, 4, 25, 0)$ است. میانگین مربعات مساوی است با $\frac{76}{9} = 8.44$ و ریشه میانگین (دامنه آرام‌اس) برابر با $2/9$ است. اما اگر شکل موجی با دقتاً همین دامنه قله اما با فرکانس بیشتر به‌صورت $(0, 5, 0, -5, 0, 5, 0, -5, 0)$ داشته باشیم، دامنه آرام‌اس آن بالاتر $(\sqrt{100/9} = 3.33)$ خواهد بود. علاوه بر آن، دامنه آرام‌اس امواج مرکبی که موج آنها شکل‌های متفاوتی دارد، نیز ممکن است متفاوت باشد، اگرچه دامنه قله آنها تفاوتی نداشته باشد. به‌عنوان مثال، موجی با نمونه‌های $(0, 0, 0, -5, 0, 0, 0, 0, 5, 0, 0, 0)$ ، دامنه آرام‌اس پایین‌تری $(2/36)$ نسبت به دامنه آرام‌اس مثال اول دارد، هرچند که فرکانس آنها یکسان است.

دامنه آرام‌اس در بسیاری از نرم‌افزارهای تجزیه گفتار برای رسم «نمودار دامنه» به کار می‌رود. این نمودار، اندازه‌های دامنه را که از پنجره‌های متوالی یا هم‌پوشان در سیگنال گفتار گرفته شده‌اند، نشان می‌دهد.

هنگام انتخاب اندازه پنجره آرام‌اس دو ملاحظه عملی را باید در نظر داشت (البته اگر برنامه به شما اجازه تغییر اندازه پنجره را بدهد). اول اینکه، هر قدر طول پنجره بلندتر باشد، منحنی دامنه حاصل هموارتر خواهد بود. مثلاً اگر طول پنجره ۲۰ میلی‌ثانیه یا بیشتر باشد، از آنجا که یک مقدار آرام‌اس برابر با میانگین مربع نمونه‌ها در طول یک پنجره است، بنابراین پالس‌های چاکنایی به‌صورت قله‌های منفرد در منحنی دامنه مشاهده نمی‌شوند. دوم اینکه، هنگامی که یک پنجره تجزیه بلند به کار

1. windowing

می‌رود، دقت زمانی منحنی دامنه کاهش پیدا می‌کند. هر قدر اندازه پنجره افزایش پیدا کند، توانایی منحنی آرام‌اس برای نشان دادن تغییرات ناگهانی در شکل موج آکوستیکی کاهش پیدا می‌کند. مثلاً اگر مدت زمان یک پنجره ۵ میلی‌ثانیه باشد، اما طول پنجره ۲۰ میلی‌ثانیه و گام بین پنجره‌های متوالی ۱۰ میلی‌ثانیه باشد، مدت زمان دیرش پنجره در منحنی آرام‌اس چقدر خواهد بود؟ به‌طور کلی نباید مقادیر دیرش را با منحنی دامنه آرام‌اس اندازه‌گیری کنید.

۳-۲. تبدیل سریع فوریه^۲

تبدیل سریع فوریه (افاف‌تی)، یک روش پردازش دیجیتالی سیگنال برای محاسبه تبدیل فوریه (طیف) یک سیگنال است (کولی و دیگران،^۳ ۱۹۶۹). اگر از فصل یک به خاطر داشته باشید، تجزیه فوریه یک شکل موج آکوستیکی را تبدیل به یک طیف می‌کند و مؤلفه‌های سینوسی آن را نشان می‌دهد (شکل‌های ۱-۵، ۱-۶ و ۱-۷). افاف‌تی روشی برای محاسبه تبدیل فوریه با استفاده از کامپیوتر است. مهم‌ترین نکته‌های عملی که هنگام استفاده از الگوریتم افاف‌تی برای تجزیه فوریه باید در نظر داشت عبارت‌اند از: وضوح زمانی^۴ و وضوح فرکانسی^۵. طبیعی است که بین این دو، یک رابطه بده‌بستان وجود دارد.

در الگوریتم افاف‌تی این رابطه بده‌بستان ناشی از رابطه‌ای است که بین اندازه پنجره و وضوح فرکانسی وجود دارد. طیف فرکانسی از صفر هرتز تا فرکانس نایکویست (نصف نرخ نمونه‌برداری) به وسیله تعدادی نقطه با فاصله مساوی از یکدیگر بازنمایی می‌شود و تعداد نقاط بر اساس تعداد نمونه‌ها در پنجره شکل موج تعیین می‌شود (به مطلب «پنجره‌گذاری» رجوع کنید). علاوه بر آن، برای راحتی و سرعت عمل در محاسبه (مفهوم «سریع» در عبارت «تبدیل سریع فوریه»)، تعداد نمونه‌ها در پنجره افاف‌تی باید توانی از ۲ باشد (مثلاً $2^7 = 128$ ، $2^8 = 256$ ، $2^9 = 512$ و غیره).

1. click

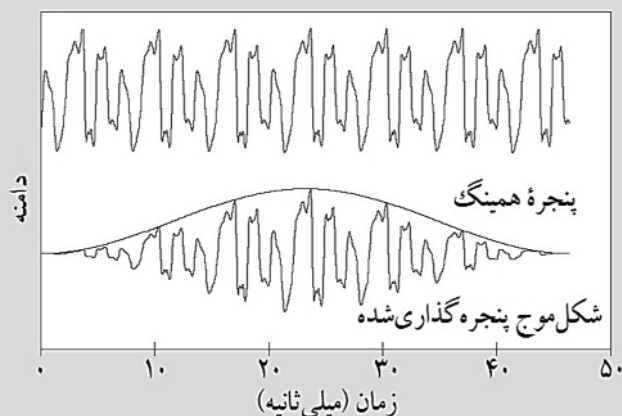
2. fast Fourier transform (FFT)

3. Cooley et al.

4. time resolution

5. frequency resolution

پنجره گذاری



شکل ۸-۳ «پنجره گذاری» با استفاده از پنجره همینگ. در بالا یک قطعه شکل موج وجود دارد و در پایین، یک پنجره همینگ هست که پوش^۱ شکل موج «پنجره گذاری شده» را نشان می دهد.

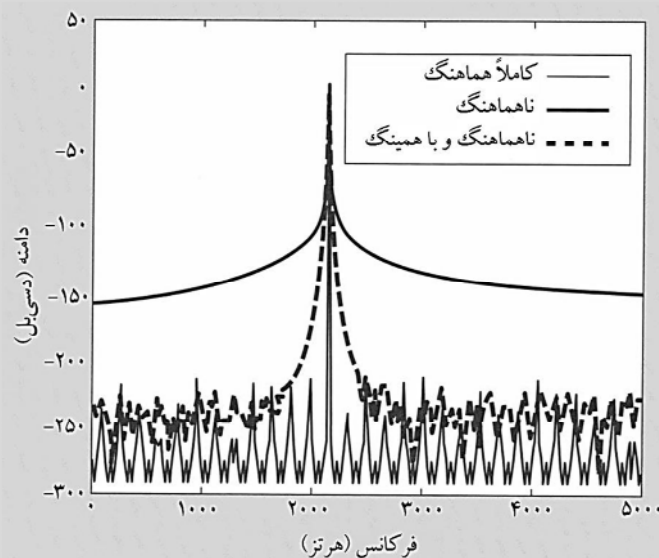
یک پنجره شکل موج، یک بخش یا قطعه ای از شکل موج است که پنجره تجزیه روی آن قرار گرفته است. در فصل یک گفتیم که سیگنال گذرا طیفی مسطح دارد (شکل های ۱۰-۱ و ۱۱-۱). ارتباط این موضوع با تجزیه افافتی در این است که اگر قطعه ای اختیاری از یک شکل موج دیجیتال را انتخاب کنید، معمولاً دامنه انرژی اولین و آخرین نمونه در آن قطعه صفر نیست و در واقع، موج گذراست. در نتیجه، تجزیه طیفی قطعه شکل موج به اشتباه همراه با طیف یک سیگنال گذرا خواهد بود. پنجره گذاری این مسئله را با تغییر مقادیر دامنه شکل موج حل می کند، به نحوی که دامنه نمونه های نزدیک به لبه های قطعه، پایین و دامنه نمونه های میانی قطعه، کاملاً بالا خواهد بود. شکل ۸-۳ کاربرد یک پنجره را نشان می دهد. در برنامه های کامپیوتری اغلب امکان انتخاب پنجره های مختلفی وجود دارد. دو تا از رایج ترین آنها، پنجره های «همینگ»^۲ و «مستطیلی» نامیده می شوند. همان طور که شکل ۸-۳ نیز نشان می دهد، پنجره همینگ دامنه نمونه ها را در کناره لبه های قطعه موج کاهش می دهد، اما پنجره مستطیلی نمونه های

←

1. envelope

2. hamming

شکل موج را اصولاً تغییر نمی‌دهد. پنجره همینگ برای تجزیه افافتی و پنجره مستطیلی برای سایر روش‌های تجزیه‌ای که بعداً در این فصل مطرح می‌شوند، مثل محاسبه زیربمی با روش خودهمبستگی^۱ و تجزیه ال‌پی‌سی^۲ به کار می‌رود.



شکل ۳-۹. «پنجره‌گذاری» تجزیه افافتی را بهبود می‌بخشد. اگر دوره تناوب سیگنال کاملاً به اندازه طول پنجره تجزیه باشد، پنجره همینگ لازم نیست. اما هنگامی که سیکل‌های شکل موج کاملاً قالب پنجره تجزیه نیستند، کاربرد پنجره همینگ (یا شبیه آن) قبل از تجزیه افافتی ضروری است.

مثال زیر نشان می‌دهد که چگونه پنجره‌گذاری سبب بهبود تجزیه افافتی می‌شود. شکل ۳-۹ سه طیف مختلف افافتی را نشان می‌دهد. اولین طیف (خط پیوسته نازک با تعداد زیادی قله‌های تیز کم‌انرژی)، طیف یک موج سینوسی است که دوره تناوب آن کاملاً به اندازه تعداد نمونه‌ها در پنجره افافتی است. یعنی طول پنجره دقیقاً مضربی از دوره تناوب موج سینوسی است. با نگاهی به این طیف می‌بینید که یک قله تیز در فرکانس موج سینوسی (اندکی بیش از ۲۰۰۰ هرتز) در طیف وجود دارد و مابقی طیف، دامنه خیلی پایینی، بیش از ۲۰۰ دسی‌بل کمتر از قله طیف دارد. طیف دوم که با خط پیوسته ضخیم رسم شده (تابعی هموار با کمترین مقدار دامنه در حدود ۱۶۰- دسی‌بل)، طیف یک موج سینوسی است که

1. auto-correlation pitch tracking

2. LPC

فرکانس آن اندکی متفاوت است و بنابراین، دقیقاً قالب پنجره تجزیه افافتی نیست. همان‌طور که می‌بینید این ناهماهنگی جزئی بین دیرش دوره تناوب با طول پنجره تجزیه منجر به ایجاد طیفی شده است که در آن دامنه قله موج سینوسی اختلاف کمی با دامنه «سکوت» در پس‌زمینه دارد. در نهایت، طیف سوم که با خط چین پررنگ رسم شده، نتیجه اعمال پنجره همینگ بر شکل موج ناهماهنگ است. طیف حاصل بسیار شبیه به طیفی است که وقتی شکل موج دقیقاً قالب پنجره تجزیه باشد، تولید می‌شود [طیف اول]. برای یک سیگنال مرکب مثل گفتار، پنجره‌گذاری در تجزیه افافتی ضروری است.

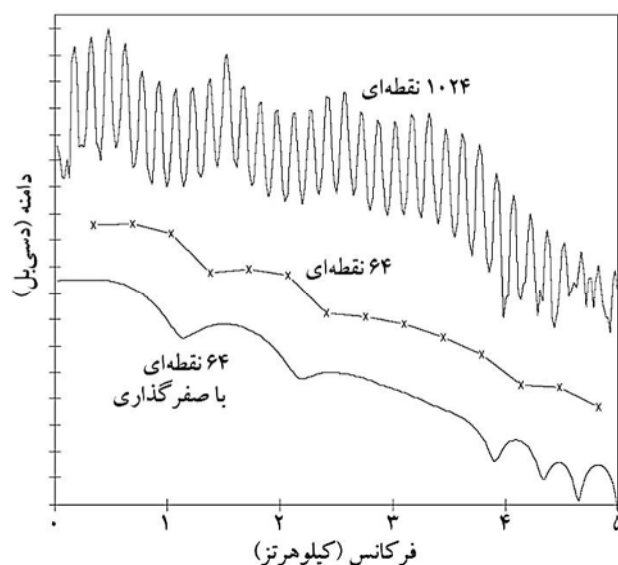
به عنوان مثال، اگر نرخ نمونه‌برداری ۲۲ کیلوهرتز و اندازه پنجره افافتی ۱۰۲۴ نمونه باشد، که برابر با یک قطعه ۴۶/۵ میلی‌ثانیه‌ای از شکل موج است (۱۰۲۴ نمونه تقسیم بر ۲۲۰۰۰ نمونه در ثانیه برابر است با ۰/۰۴۶۵ ثانیه) فاصله بین نقاط در طیف حاصل برابر با ۲۱/۴۸ هرتز است ($\frac{22000}{1024}$) و بنابراین، برآوردی از دامنه طیفی در ۲۱/۴۸ هرتز، ۴۲/۹۷ هرتز، ۶۴/۴۵ هرتز و الی آخر به دست می‌آید. الگوریتم افافتی دامنه طیفی را برای نقاط مابین این فرکانس‌ها محاسبه نمی‌کند، اما اگر فاصله بین نقاط متوالی در طیف کم باشد، می‌توان مقادیر آنها را با اتصال^۱ نقاط به یکدیگر تخمین زد (همان‌طور که یک برنامه کامپیوتری در یک نمودار طیفی «نقاط را به یکدیگر وصل می‌کند»).

شکل ۳-۱۰ تفاوت بین طیف‌های افافتی یک واکه را که با یک پنجره تجزیه بلند (طیف بالا) و یک پنجره تجزیه کوتاه (طیف میانی) محاسبه شده‌اند، نشان می‌دهد. دو نکته قابل توجه در مورد این طیف‌ها وجود دارد. اول اینکه، هنگامی که پنجره تجزیه بلند است، جزئیات طیف بهتر تفکیک می‌شود. طیف با جزئیات بیشتر که از پنجره تجزیه بلند (۱۰۲۴ نمونه) به دست آمده، قله‌های باریک با فاصله مساوی یعنی همسازها را نشان می‌دهد، اما طیفی که فقط با ۶۴ نمونه از همان شکل موج به دست آمده، تنها شکل کلی طیف را نشان می‌دهد و از آنجا که تعداد نقاط طیفی کم است، اعوجاج^۲ دارد. دوم اینکه، برای به دست آوردن وضوح فرکانسی بیشتر، اندازه

1. interpolation

2. distorted

پنجره تجزیه باید در حدود ۴۶/۵ میلی ثانیه [۱۰۲۴ نمونه] باشد. اما افزایش طول پنجره مسئله ساز است، زیرا در تجزیه گفتار، موارد بسیاری پیش می آید که طیف در مدت زمان بسیار کوتاهی (۵ تا ۱۰ میلی ثانیه) تغییر می کند و از این رو، یا این ویژگی های زمانی در طیف اف اف تی منتج از یک پنجره بلند ذخیره نمی شود و یا در حالت بدتر، یک پنجره بلند مخلوطی از صداها را دربرمی گیرد و بنابراین، طیف آن، ویژگی صداها را مجزا اما مجاور هم را خواهد داشت. با استفاده از یک پنجره تجزیه کوتاه تر این مشکل رفع خواهد شد. با نرخ نمونه برداری ۲۲ کیلوهرتز اندازه پنجره ۶۴ نقطه ای، ۳ میلی ثانیه است، اما با این پنجره جزئیات طیف ذخیره نمی شود و شکل دقیق طیف بسیار وابسته به مکان پنجره تجزیه است.



شکل ۳-۱۰. اندازه پنجره و وضوح فرکانسی در تجزیه اف اف تی. طیف بالا که با عنوان «نقطه ای ۱۰۲۴» مشخص شده، با استفاده از یک پنجره تجزیه ۴۶ میلی ثانیه ای تولید شده است. بنابراین، وضوح فرکانسی آن بالا و وضوح زمانی آن پایین است. طیف میانی با عنوان «نقطه ای ۶۴» با استفاده از یک پنجره تجزیه ۳ میلی ثانیه ای تولید شده است. بنابراین، وضوح زمانی بالا و وضوح طیفی پایین دارد. در این طیف، هر یک از مقادیر طیفی با علامت X مشخص شده است. طیف پایین که با عنوان «نقطه ای ۶۴ با صفر گذاری» مشخص شده، با تجزیه همان قطعه ۳ میلی ثانیه ای موج اما به وسیله یک پنجره تجزیه ۴۶ میلی ثانیه ای که صفر گذاری شده، تولید شده است. فاصله بین مقادیر دامنه بر روی محور عمودی ۲۰ دسی بل است.

پایین‌ترین طیف در شکل ۳-۱۰ ترکیب جالبی از پنجره تجزیه کوتاه و بلند را نشان می‌دهد. این طیف، حاصل تجزیه ۶۴ نمونه در یک پنجره افافتی ۱۰۲۴ نقطه‌ای است به گونه‌ای که مقدار ۹۶۰ نمونه باقیمانده در پنجره افافتی، صفر شده است (این عمل صفرگذاری^۱ نامیده می‌شود). یعنی ما طول پنجره تجزیه و در نتیجه، وضوح فرکانسی را افزایش می‌دهیم بی‌آنکه تعداد نمونه‌های مورد تجزیه در سیگنال افزایش پیدا کند. طیف حاصل شبیه به طیف ۶۴ نقطه‌ای است، از این جهت که فقط نمای کلی طیف را نشان می‌دهد، اما با این تفاوت که تعداد نقاط طیفی دامنه زیاد است. بنابراین، در حالی که طیف ۶۴ نقطه‌ای (میانی) ممکن است دندان‌دندانه باشد، طیف ۶۴ نقطه‌ای صفرگذاری شده (پایینی)، نمای کلی طیف را دقیق‌تر نشان می‌دهد، زیرا فاصله فرکانسی بین نقاط طیفی به همان میزانی است که در طیف ۱۰۲۴ نقطه‌ای است.

صفرگذاری امکان تعیین تعداد اختیاری نمونه برای تجزیه طیفی، بدون نگرانی از میزان فاصله فرکانسی بین نقاط طیفی میسر می‌سازد و در نتیجه، امکان تعیین میزان اختیاری وضوح زمانی و محوی طیفی^۲ را میسر می‌کند. پنجره‌های کوتاه بسیار حساس به جایی هستند که بر روی شکل موج قرار می‌گیرند و پنجره‌های بلند، تغییرات سریع طیفی را نشان نمی‌دهند، اما حداقل تعداد نمونه‌ها در طیف محاسبه شده همواره مناسب است. از این روش در طیف‌نگارهای دیجیتال برای کنترل میزان محوی طیفی («پهنای فیلتر تجزیه») در تولید طیف‌نگاشت‌های گفتاری استفاده می‌شود (برای مطالب بیشتر بخش ۳-۳-۶ را ببینید).

۳-۳-۳. محاسبه زیرویمی با روش خودهمبستگی

به نظر می‌رسد که پالس‌های متوالی زیرویمی در گفتار واکدار شبیه به هم هستند (به عنوان مثال شکل ۳-۱۱ را ببینید). از این رو، اگر بخواهیم زیرویمی واک را اندازه بگیریم، تنها کاری که باید انجام دهیم این است که به الگوی تکراری قطعات شبیه به هم شکل موج توجه نماییم و سپس، دیرش هر تکرار را اندازه‌گیری کنیم. به لحاظ

1. zero padding

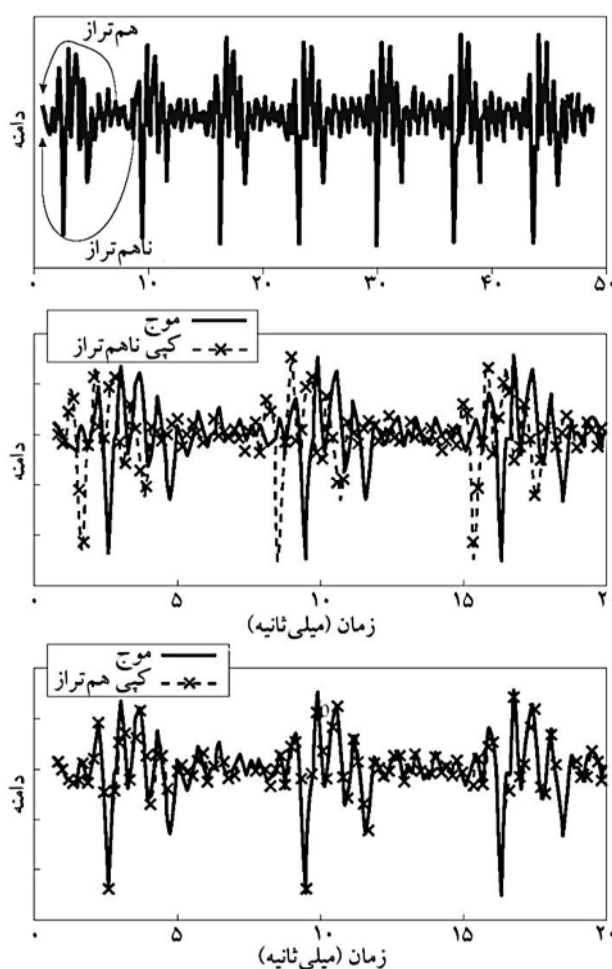
2. spectral smearing

ریاضی، این روش مستلزم یافتن قطعاتی از شکل موج است که با یکدیگر همبستگی بالایی دارند. به این روش خودهمبستگی گفته می‌شود، زیرا قطعات شکل موج متعلق به یک سیگنال صوتی هستند. روش خودهمبستگی را می‌توان برای یافتن الگوهای تکراری حتی در درون یک پالس زیروبمی نیز به کار برد و در واقع در «کدگذاری پیشگویانه خطی»^۱ (الپی‌سی) از این روش برای یافتن قله‌ها در طیف گفتار استفاده می‌شود (رجوع کنید به بخش ۳-۳-۵).

شکل ۳-۱۱ محاسبه زیروبمی با روش خودهمبستگی را نشان می‌دهد. اگر سیکل‌های متوالی یک موج مرکب را روی یکدیگر قرار دهید، خواهید دید که با یکدیگر همبستگی بسیار بالایی دارند، یعنی شکل موج پس از گذشت یک دوره تناوب زیروبمی^۲ با خودش همبستگی دارد. یک کپی از شکل موج از نقطه‌ای که در شکل ۳-۱۱ با عنوان «ناهم‌تراز» مشخص شده، بردارید و آن را روی نقطه شروع موج، همان‌طور که فلش نشان می‌دهد، بگذارید. خواهید دید که دو قطعه بر یکدیگر منطبق نمی‌شوند (همان‌طور که در شکل وسط می‌بینید). اما اگر هم‌ترازی به‌درستی صورت گیرد، به‌گونه‌ای که کپی شکل موج دقیقاً از یک سیکل زیروبمی بعد از شروع موج برداشته شود (در شکل با عبارت «هم‌تراز» مشخص شده) در این صورت، هر دو قطعه شکل موج تقریباً بر یکدیگر منطبق خواهند شد (شکل پایین). روش خودهمبستگی از این ویژگی صوت واک‌دار برای محاسبه خودکار فرکانس پایه (F_0) با یافتن دیرش دوره‌های تناوب استفاده می‌کند. «نمودار زیروبمی»، مقدار F_0 ارتعاش تارهای صوتی را به‌صورت تابعی در زمان نشان می‌دهد، بدین‌صورت که محاسبه مقادیر F_0 در فواصل زمانی ۱۰ یا ۲۰ میلی‌ثانیه صورت می‌گیرد. در روش خودهمبستگی، قطعه کوچکی از شکل موج آکوستیکی (یک پنجره شکل موج) که معمولاً شامل چندین سیکل دوره تناوب است انتخاب شده، سپس محاسبه همبستگی برای تعدادی از طول‌های ممکن دوره تناوب صورت می‌گیرد. طول دوره تناوبی که بیشترین میزان همبستگی را نشان دهد، گزارش می‌شود.

1. linear predictive coding

2. pitch period



شکل ۳-۱۱. سیکل‌های متوالی زیروبمی در آوای واک‌دار با یکدیگر همبستگی دارند. شکل بالا ۵۰ میلی‌ثانیه از واکه چروکی^۱ [ʌ] را نشان می‌دهد. شکل وسط نتیجه رسم شکل موج اصلی با یک کپی ناهم‌تراز آن را نشان می‌دهد. شکل پایین نتیجه رسم شکل موج اصلی با کپی‌ای را نشان می‌دهد که دقیقاً یک سیکل زیروبمی بعد از موج اصلی آغاز می‌شود.

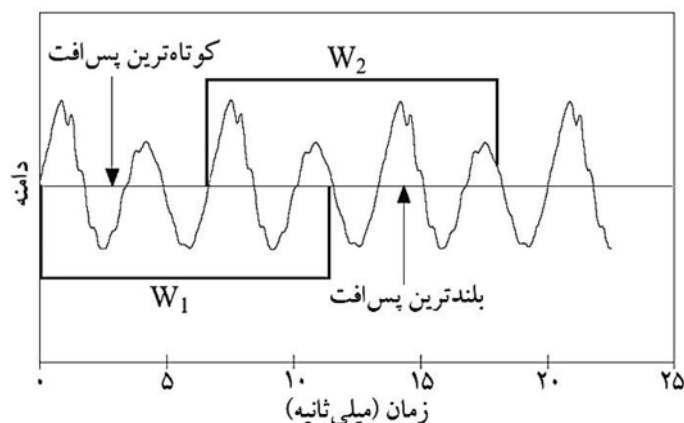
شکل ۳-۱۲، شیوه دیگری از روش خودهمبستگی برای محاسبه زیروبمی را نشان می‌دهد. این شکل دو پنجره هم‌پوشان (دو قطعه شکل موج) در یک موج مرکب متناوب

۱. Cherokee. یکی از قبایل بومی آمریکا. (مترجم)

را نشان می‌دهد. امواج مرکب در این دو پنجره همبستگی بالایی با یکدیگر دارند، زیرا فاصله بین شروع W_1 تا شروع W_2 یعنی پس‌افت^۱ خودهمبستگی دقیقاً برابر با دیرش یک سیکل است. در این روش برای به‌دست‌آوردن زیروبمی موج، همبستگی بین دو پنجره W_1 و W_2 در پس‌افت‌های متعدد بین نقاط مشخص شده با عنوان «کوتاه‌ترین پس‌افت» و «بلندترین پس‌افت» محاسبه می‌شود. بدین‌صورت که ابتدا پنجره W_2 در آغاز «کوتاه‌ترین پس‌افت» قرار می‌گیرد و همبستگی بین W_1 و W_2 محاسبه می‌شود. سپس پنجره W_2 به نمونه سمت راست در شکل موج دیجیتال حرکت می‌کند و دوباره همبستگی محاسبه می‌شود. حرکت پنجره W_2 به سمت راست ادامه پیدا می‌کند تا زمانی که نقطه شروع پنجره برابر با «بلندترین پس‌افت» باشد. طول پس‌افتی که بیشترین میزان همبستگی بین دو پنجره W_1 و W_2 را نشان دهد، به‌عنوان مدت‌زمان یک دوره تناوب موج در نظر گرفته می‌شود. معکوس دوره تناوب نیز برابر با فرکانس پایه موج مرکب است (اگر T برابر باشد با مدت‌زمان یک دوره تناوب برحسب ثانیه آنگاه F_0 برحسب Hz برابر است با $1/T$).

یک راه برای افزایش سرعت روش خودهمبستگی، محدودکردن تعداد سیکل‌های مورد محاسبه است. بدین‌طریق، تعداد همبستگی‌های مورد محاسبه محدود می‌شود. بسیاری از برنامه‌های تجزیه گفتار این کار را یا با تخمین میانگین F_0 گویشور یا در محدوده‌ای از مقادیر ممکن F_0 انجام می‌دهند. اینها پارامترهایی هستند که کاربر می‌تواند آنها را تغییر دهد و در واقع، در بسیاری موارد برای محاسبه دقیق زیروبمی گفتار گویشوران خاص باید آنها را تغییر دهد. دو خطای رایج روش‌های محاسبه زیروبمی با خودهمبستگی عبارت‌اند از: نصف‌شدن زیروبمی و دوبرابرشدن زیروبمی. هر دوی این خطاها ممکن است روی دهد، حتی اگر محدوده F_0 پیش‌بینی شده به‌وسیله برنامه درست باشد. اما اگر پارامترهای برنامه با محدوده زیروبمی گویشور هماهنگی نداشته باشد، بی‌تردید یکی از این خطاها روی می‌دهد.

1. lag



شکل ۳-۱۲. نمایش روش خودهمبستگی برای محاسبهٔ زیروبمی. پنجرهٔ دوم (W_2) در «پس‌افت‌های» مختلف، از «کوتاه‌ترین پس‌افت» تا «بلندترین پس‌افت»، شروع به حرکت از پنجرهٔ اول (W_1) می‌کند. دیرش دورهٔ تناوب برابر با دیرش پس‌افتی است که بالاترین همبستگی بین قطعهٔ شکل موج در W_1 با قطعهٔ شکل موج در W_2 را نشان دهد.

نصف شدن زیروبمی زمانی روی می‌دهد که دو سیکل زیروبمی در درون پنجره‌های W_1 و W_2 قرار می‌گیرد (شکل ۳-۱۲) و «کوتاه‌ترین پس‌افت» پس از پایان اولین سیکل زیروبمی پنجرهٔ W_1 است. در این حالت، برنامهٔ محاسبهٔ زیروبمی فرض می‌کند که دورهٔ تناوب زیروبمی بلندتر از میزانی است که عملاً در سیگنال روی می‌دهد. در نتیجه، بهترین همبستگی را وقتی پیدا می‌کند که پنجرهٔ W_2 در آغاز سومین سیکل زیروبمی قرار می‌گیرد (به‌جای حالت عادی که بهترین همبستگی، مثل شکل ۳-۱۲، موقعی یافت می‌شود که W_2 در آغاز دومین سیکل زیروبمی قرار می‌گیرد). هنگامی که W_2 در آغاز سومین سیکل قرار می‌گیرد، مقدار پس‌افت برابر است با دیرش دو دورهٔ تناوب زیروبمی. از این رو، مقدار F_0 گزارش‌شده نصف مقدار F_0 واقعی است. نصف‌شدن زیروبمی حتی وقتی که محدودهٔ F_0 پیش‌بینی‌شده درست باشد، نیز می‌تواند روی دهد. این وضعیت معمولاً وقتی روی می‌دهد که سیکل‌های متناوب زیروبمی شباهت بیشتری به یکدیگر دارند تا سیکل‌های مجاور، همان‌طور که در بعضی انواع واک‌سازی مثل جز جز چاکنایی^۱ و یا دوصدایی^۲ دیده می‌شود.

1. vocal fry

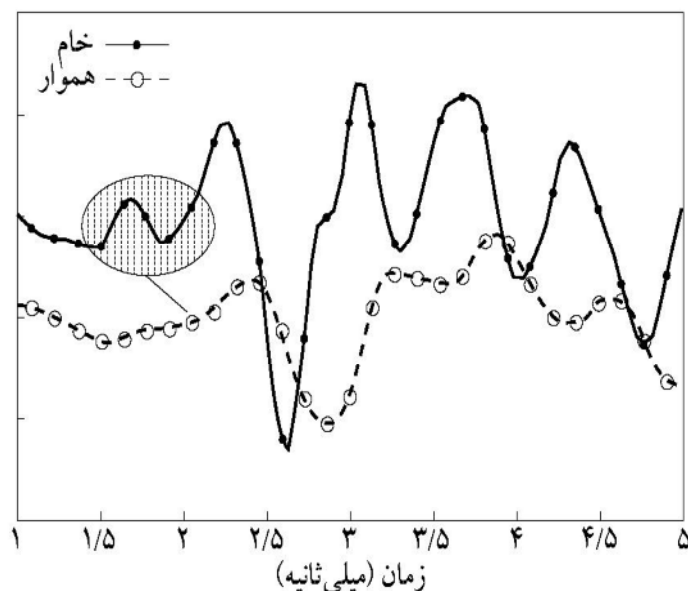
۲. diplophonia. در این نوع واک‌سازی، دو واک با دو زیروبمی تولید می‌شود. (مترجم)

دوبرابر شدن زیروبمی زمانی روی می‌دهد که کوتاه‌ترین پس‌افت مورد محاسبه به کوتاهی نصف یک سیکل دوره تناوب باشد و نیمه دوم سیکل بسیار شبیه به نیمه اول آن باشد. همان‌طور که شکل ۳-۱۲ نیز این را نشان می‌دهد، هر سیکل دوره تناوب شکل موج را می‌توان به دو نیمه بسیار شبیه به هم تقسیم نمود؛ هر سیکل دوره تناوب از یک قله مثبت و به دنبال آن یک قله منفی و سپس یک قله مثبت دیگر و به دنبال آن یک قله منفی دیگر تشکیل یافته است. برنامه محاسبه خودهمبستگی به‌سادگی ممکن است فرض کند که دو نیمه یک سیکل، سیکل‌های مجزایی هستند و اگر «بلندترین پس‌افت» کمتر از یک سیکل کامل دوره تناوب باشد، بدون استثنا این محاسبه اشتباه روی خواهد داد. اگر نصف دوره تناوب به اشتباه به عنوان دیرش کامل دوره تناوب فرض شود، مقدار F_0 گزارش شده دوبرابر مقدار واقعی F_0 است.

۳-۴. صافی‌های دیجیتال

سیگنال‌های صوتی دیجیتال را به‌آسانی می‌توان ویرایش کرد و تغییر داد. با برخی عملیات ویرایش در برنامه وُرد^۱ آشنا هستید. مثلاً ویرایشگرهای موج به شما امکان کپی، برش و انتقال قطعه‌های شکل موج را می‌دهند. همچنین، می‌توان دامنه یک شکل موج دیجیتال را با ضرب مقدار هر نمونه در یک ضریب تغییر داد. به عنوان مثال، برای کاهش دامنه به نصف، آن را ضربدر ۰/۵ و یا برای دو برابر کردن، آن را ضربدر ۲ می‌کنیم. حتی اندکی پیچیده‌تر، می‌توانیم صافی‌های صوتی دیجیتال را برای فیلتر پایین‌گذر، بالاگذر و نوارگذر بسازیم. در این فصل مبحث صافی‌های دیجیتال به‌طور کامل مورد بحث قرار نمی‌گیرد (در این مورد کتاب لاینز (۱۹۹۷) را پیشنهاد می‌کنم)، اما امیدوارم با معرفی برخی مفاهیم پایه، چگونگی استفاده از نرم‌افزار صافی دیجیتال را یاد گرفته و درک مفهوم تجزیه ال‌پی‌سی (که بعداً در این فصل مطرح خواهد شد) برایتان آسان‌تر شود.

با یک روش رایج در هموارسازی داده‌ها به نام «هموارسازی میانگین متحرک»^۱ شروع می‌کنیم. بی‌تردید میانگین‌گرفتن از مجموعه‌ای از رقم‌ها سبب جایگزین‌شدن یک مقدار به جای مجموعه مقادیر می‌شود. از این رو، برای هموارکردن یک رشته از داده‌های زمانی (مثلاً یک شکل موج صوتی) می‌توان مقدار هر نمونه را با میانگین نمونه‌های مجاور جایگزین کرد. نمونه‌ای از هموارسازی با روش میانگین متحرک در شکل ۳-۱۳ نشان داده شده است.



شکل ۳-۱۳. از نقاط یک موج صوتی خام برای تولید یک موج هموار، میانگین گرفته می‌شود. موج هموار حرکت آرام و فرکانس پایین شکل موج خام را در خود حفظ می‌کند، اما فاقد مؤلفه‌های لرزان^۲ با فرکانس بالا و حرکت تند شبیه نوسان درون بیضی است.

تکنیک هموارسازی با روش میانگین متحرک، نمونه ساده‌ای از یک صافی پایین‌گذر است، زیرا تغییرات سریع از یک نمونه به نمونه دیگر در شکل موج خام حذف (فیلتر) می‌شوند. یکی از مهم‌ترین نکاتی که باید در رابطه با یک صافی بدانیم، منحنی پاسخ فرکانسی آن است. برای به‌دست‌آوردن منحنی پاسخ فرکانسی صافی

1. moving average smoothing 2. jittery

میانگین متحرک باید صافی را با فرمولی تعریف کنیم که ضرایب صافی را جدا کند. فرض کنید که شکل موج هموار شده را $y(n)$ و شکل موج خام را $x(n)$ بنامیم. n به یک نقطه در زمان، زمان n اشاره دارد و x و y اسم‌های اختیاری هستند که به شکل موج‌های دیجیتال نمونه‌برداری شده می‌دهیم. اکنون با این علائم اختصاری می‌توان معادله (۱-۳) را برای به‌دست آوردن شکل موج هموار از شکل موج خام به‌صورت زیر تعریف کرد:

$$Y(n) = \frac{x(n-4) + x(n-3) + x(n-2) + x(n-1) + x(n-0)}{5} \quad (1-3)$$

مقدار موج هموار شده $y(n)$ در زمان n برابر است با میانگین پنج نمونه قبلی $x(n-4) \dots x(n-0)$ در شکل موج خام. می‌توان معادله فوق را به‌صورت معادله (۲-۳) ساده کرد:

$$\begin{aligned} Y(n) &= \frac{1}{5}x(n-4) + \frac{1}{5}x(n-3) + \frac{1}{5}x(n-2) + \frac{1}{5}x(n-1) + \frac{1}{5}x(n-0) \\ &= \sum_{k=0}^{M-1} \frac{1}{5} x(n-k) \\ &= \sum_{k=0}^{M-1} h_k x(n-k) \end{aligned} \quad (2-3)$$

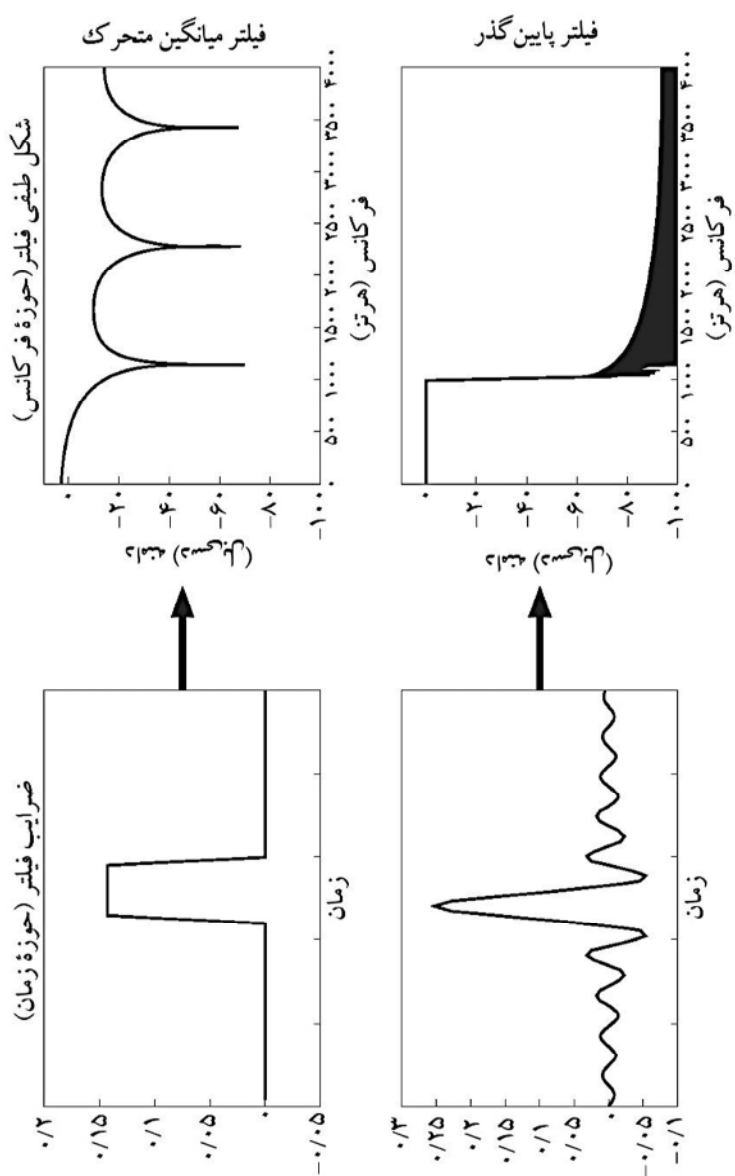
ما ابتدا مقدار $\frac{1}{5}$ را در همه نمونه‌های شکل موج ضرب کردیم، سپس فرمول را به‌صورت یک حاصل جمع با علامت سیگمای بزرگ $\sum$ نوشتیم [در این فرمول $x(n-k)$ نیز جایگزین نمونه‌های شکل موج خام $x(n-4) \dots x(n-0)$ شده است]؛ یعنی ضرب $\frac{1}{5}$ در همه مقادیر موج خام به‌ازای k از صفر تا $M-1$ با یکدیگر جمع می‌شود. M تعداد نمونه‌هایی است که میانگین آنها محاسبه می‌شود (در این مثال M برابر ۵ نمونه است)^۱ و در نهایت، می‌توان مقدار ثابت $\frac{1}{5}$ را با متغیر جدید $h(k)$ جایگزین کرد که

۱. بنابراین، در این مثال مقدار k از صفر تا ۴ است. (مترجم)

در این مثال به صورت $h(0)=\frac{1}{5}$, $h(1)=\frac{1}{5}$, ..., $h(4)=\frac{1}{5}$ خواهد بود. متغیر جدید $h(k)$ در واقع، صافی میانگین متحرک است، به گونه‌ای که مقادیر آن ضرایب صافی محسوب می‌شوند و ویژگی‌های پاسخ فرکانسی صافی را تعیین می‌کنند. برای تعیین پاسخ فرکانسی یک صافی تنها کاری که لازم است انجام دهیم این است که تجزیه افافتی را بر روی خود ضرایب صافی به کار ببریم (به همراه صفرگذاری). نمودار بالا سمت راست در شکل ۳-۱۴، نتایج این کار را با ضرایب یک صافی میانگین متحرک هفت نقطه‌ای نشان می‌دهد. آنچه بسیار واضح است این است که هرچند هموارکننده میانگین متحرک، موج را تاحدی فیلتر پایین‌گذر می‌کند، اما به هر حال، می‌توان یک صافی پایین‌گذر خیلی کاراتر نسبت به آن تصور کرد.

برای ساختن یک صافی پایین‌گذر ایده‌آل می‌توان تبدیل فوریه را در جهت معکوس به کار برد، به این ترتیب که ابتدا شکل صافی پایین‌گذر دلخواه تعیین شده و سپس ضرایب صافی در $h(k)$ به دست می‌آید، به نحوی که اگر این ضرایب در معادله (۳-۲) قرار بگیرند، موج را به میزان مورد نظر فیلتر می‌کنند. فرایند «اعمال تبدیل فوریه در جهت عکس» تبدیل معکوس گسسته فوریه^۱ یا معکوس (دی‌اف‌تی) نامیده می‌شود. ضرایب صافی‌ای که به این طریق به دست می‌آیند، در نمودار پایین سمت چپ در شکل ۳-۱۴ نشان داده شده‌اند. صافی «ایده‌آلی» که در این شکل نشان داده شده، یک گذر تند در ۱۰۰۰ هرتز بین نوار گذر و نوار دفع دارد. همان‌طور که طیف افافتی ضرایب این صافی (نمودار پایین سمت راست) نشان می‌دهد، این صافی پایین‌گذر در واقع همه فرکانس‌های بالای ۱۰۰۰ هرتز را حذف می‌کند.

1. inverse discrete Fourier transform



شکل ۱۴-۳. مقایسه ضرایب صافی $h(k)$ با شکل طیفی فیلتر حاصل در یک صافی پایین گذر میانگین متحرک و یک صافی پایین گذر ایده آل که طراحی آن به گونه ای است که یک گذر بی نهایت تند بین نوار عبور (زیر ۱۰۰۰ هرتز) و نوار دفع (بالای ۱۰۰۰ هرتز) دارد. یک تفاوت مهم بین دو صافی این است که صافی میانگین متحرک فقط هفت ضریب دارد، در حالی که صافی پایین گذر ایده آل ۵۷۹ تا.

در طراحی صافی‌های دیجیتال، برخی «شگردهای تجاری» به کار می‌رود. یکی از موارد مورد علاقه من این است که شما می‌توانید یک صافی پایین‌گذر را با ضرب $h(k)$ به یک صافی بالاگذر تبدیل کنید و به صورت کلی‌تر، می‌توانید فرکانس مرکزی یک صافی پایین‌گذر (که فرکانس مرکزی آن صفر هرتز است) را تغییر دهید و بنابراین، با ضرب کردن $h(k)$ در یک موج سینوسی دیجیتال با فرکانس مرکزی دلخواه، یک صافی نوارگذر بسازید. همه این موارد جالب هستند، اما نکات اصلی این بخش عبارتند از: (۱) یک صافی دیجیتال توالی از ضرایب $h(k)$ است و (۲) صافی دیجیتال با استفاده از یک صافی شبیه به یک هموارکننده میانگین متحرک در معادله (۲-۳) صورت می‌گیرد.

۳-۵. کدگذاری پیش‌گویانه خطی (الپی‌سی)

در موارد بسیاری در آواشناسی آکوستیک، علاقه‌مند به مشاهده قله‌های پهن طیفی هستیم که چندین قله همساز را در بر می‌گیرند و نه همسازهای منفرد را. همان‌طور که قبلاً گفته شد، یکی از روش‌های مشاهده این الگوهای پهن طیفی، محاسبه اف‌اف‌تی با یک پنجره تجزیه کوچک به همراه صفرگذاری است. یکی از روش‌های رایج دیگر برای مشاهده قله‌های پهن طیفی، تجزیه الپی‌سی یا «کدگذاری پیش‌گویانه خطی» نامیده می‌شود.

روش الپی‌سی توسط مهندسانی که روی سیستم‌های ارتباطی با نرخ بیت پایین کار می‌کردند، اختراع شد. هدف اولیه، حفظ فضای حافظه در سیستم‌های سنتز و بازشناسی گفتار و نیز برای مخابره رمزی صوت بود. تجزیه الپی‌سی به لحاظ فیزیکی یک مدل نسبتاً حساس گفتاری است که اگر خیلی ساده بخواهیم بگوییم، شامل یک منبع صوتی (ارتعاش تارهای صوتی) و یک صافی با چندین بازخوان است. به همین جهت، تجزیه الپی‌سی روش مناسبی برای تعیین خودکار مکان و وسعت قله‌های پهن در طیف گفتار است. همان‌طور که در بخش محاسبه زیروبمی با روش خودهمبستگی (۳-۳-۳) ذکر شد، هرگاه یک الگوی مکرر در شکل موج آکوستیکی

مشاهده شود (آن گونه که ارتعاش تارهای صوتی و بازخوان‌های دستگاه گفتار تولید می‌کنند)، می‌توان الگوی مکرر را با مشاهده همبستگی شکل موج با قطعات مؤخرتر خودش محاسبه کرد.

در بخش محاسبهٔ زیرویمی با روش خودهمبستگی گفتیم که یکی از خطاهای این روش، دوبرابرشدن زیرویمی است. این خطا زمانی روی می‌دهد که بخشی از دورهٔ تناوب به اشتباه به عنوان کل دورهٔ تناوب تعیین می‌شود. تجزیهٔ ال‌پی‌سی از این خطای روش خودهمبستگی با محدودکردن خودهمبستگی‌ها در درون یک سیکل زیرویمی استفاده می‌کند. به تعبیری، دوبرابرشدن زیرویمی نتیجهٔ دلخواه در تجزیهٔ ال‌پی‌سی است.

اما به هر حال، این تصور نادرستی است که دوبرابرشدن زیرویمی را به تجزیهٔ ال‌پی‌سی ربط دهیم، زیرا تلویحاً بدین معناست که ال‌پی‌سی فقط مؤلفه‌ای را پیدا می‌کند که فرکانس آن دوبرابر F_0 است. این تصور، نادرست است، زیرا روش خودهمبستگی می‌تواند فرکانس مؤلفه‌های متعدد یک موج مرکب را محاسبه کند.

پیش تقویت^۱

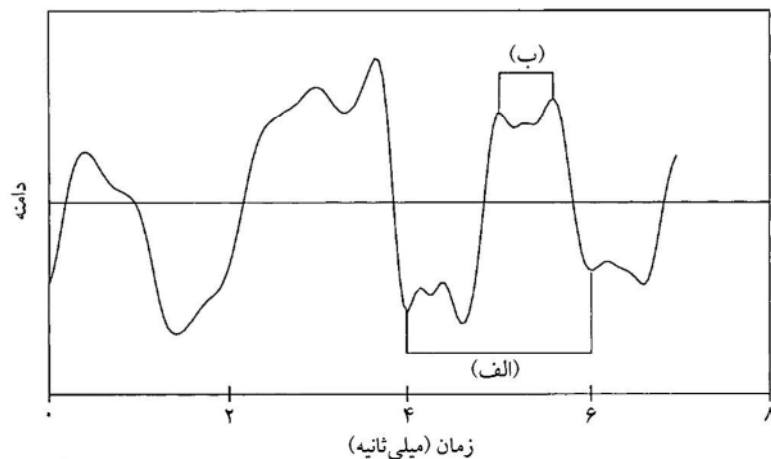
گاهی اوقات در تجزیهٔ طیفی، علاقه‌مند به افزایش دامنهٔ مؤلفه‌هایی هستیم که فرکانس بالا دارند. این کار در طیف‌نگاشت‌های آنالوگ با یک صافی بالاگذر نرم (صافی شکل‌دهندهٔ فرکانس‌های بالا) که فرکانس‌های بالای طیف را افزایش می‌دهد، صورت می‌گیرد. در پردازش دیجیتال، سیگنال، تقویت فرکانس‌های بالا با محاسبهٔ «پیش تقویتِ اولین اختلاف» صورت می‌گیرد.

برای به‌دست‌آوردن «اولین اختلاف» در یک شکل موج دیجیتال، مقادیر شکل موج را با به‌دست‌آوردن اختلاف بین نمونه‌های مجاور جایگزین کنید $(y_n = x_n - x_{n-1})$. مثلاً شکل موجی با نمونه‌های $(-6, -9, -5, -1, 3, 7, 9, 5, 0)$ را فرض کنید. مقادیر این شکل موج با محاسبهٔ اختلاف‌های متوالی به صورت

←

(۳، -۴، -۴، -۴، -۲، ۴، ۹ - ۵ = ۵ - ۰ = ۵) جایگزین می‌شود. عواقب طیفی این عمل، یعنی به دست آوردن اولین اختلاف، غیرقابل باور است: طیف با شیب شش دسی‌بل با هر بار دوبرابر شدن فرکانس (در هر اکتاو) افزایش پیدا می‌کند. پس، با دوبرابر شدن فرکانس (یک اکتاو)، دامنه طیف ۶ دسی‌بل بیشتر می‌شود.

بسیاری از بسته‌های نرم‌افزاری تجزیه گفتار به شما اجازه تعیین پیش تقویت به صورت عددی بین صفر و یک را می‌دهند. اگر شما مقدار آن را ۱ تعیین کنید، نتیجه آن پیش تقویت اولین اختلاف است که در پاراگراف بالا توضیح داده شد. اگر مقدار پیش تقویت کمتر از ۱، اما بیشتر از صفر باشد، طیف افزایش پیدا می‌کند، اما نه به میزان شیب پیش تقویت اولین اختلاف. در اینجا نیز پیش تقویت با محاسبه اختلاف بین نمونه‌های مجاور به دست می‌آید، اما با این تفاوت که مقدار نمونه‌ای که کسر می‌شود (x_{n-1}) در مقدار پیش تقویت (p) ضرب می‌شود. در این صورت، فرمول پیش تقویت به صورت $y_n = x_n - px_{n-1}$ تغییر می‌کند. مقدار پیش تقویت، p ، معمولاً به صورت عددی بین ۰/۹۰ و ۰/۹۸ تعیین می‌شود.



شکل ۳-۱۵. مؤلفه‌های فرکانسی در یک دوره تناوب زیرومی که با روش خودهمبستگی در ال‌پی‌سی تعیین شده‌اند. این شکل، یک دوره تناوب واکه شوآ را نشان می‌دهد. طول مؤلفه (الف)، ۲ میلی‌ثانیه، بنابراین، فرکانس آن ۵۰۰ هرتز و طول مؤلفه (ب)، ۰/۶۶ میلی‌ثانیه و بنابراین، فرکانس آن ۱۵۰۰ هرتز است.

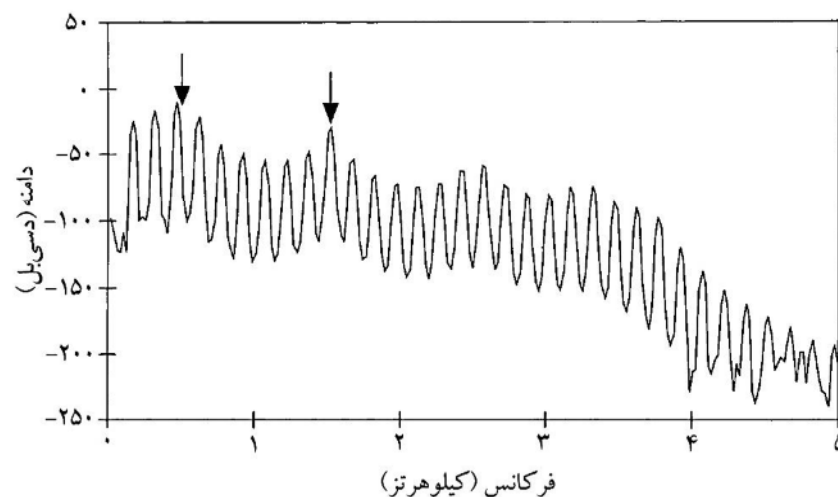
شکل ۳-۱۵ مؤلفه‌های یک سیکل زیروبمی واکه شوآ، اولین واکه کلمه ahead را نشان می‌دهد. دقت کنید که کل شکل فقط یک دوره تناوب زیروبمی واکه را نشان می‌دهد و بنابراین، «دوره تناوب» دو قطعه (الف) و (ب) که در شکل مشخص شده است، کمتر از یک دوره تناوب زیروبمی واکه است و در نتیجه، فرکانس آنها بالاتر است. این فرکانس‌ها، فرکانس قله‌های پهن طیفی را نشان می‌دهند.

همان‌طور که گفتیم روش خودهمبستگی دیرش پس‌افت‌ها را به‌گونه‌ای پیدا می‌کند که الگوی شکل موج در پایان پس‌افت، تقریباً آیینۀ تصویر خودش است. این را به راحتی می‌توان برای قطعه مشخص شده با عنوان (الف) در شکل ۳-۱۵ ملاحظه نمود. این نوسان سیکل وار با دیرش ۲ میلی ثانیه، تقریباً سه بار در یک سیکل زیروبمی تکرار می‌شود. بر روی این قطعه، نوسان (ب) که بهتر است نوسان (الف) را از آن کم کنیم، با سرعت نوسان بیشتر در حال تغییر است و بر روی قله‌های (الف) موج‌های کوچکی تولید می‌کند. دیرش دوره تناوب این قطعه ۰/۶۶ میلی ثانیه است. بر اساس دیرش نوسان‌ها، فرکانس آنها نیز به دست می‌آید ($1/b = 1500 \text{ Hz}$ ، $1/a = 500 \text{ Hz}$).

شکل ۳-۱۶، طیف فوریۀ حاصل از شکل موج ۳-۱۵ را نشان می‌دهد. فلش‌ها به فرکانس‌هایی اشاره دارد که از تجزیۀ ال‌پی‌سی با روش خودهمبستگی در شکل ۳-۱۵ به دست آمده‌اند. ملاحظه کنید که طیف از همسازهای فرکانس پایه تشکیل یافته است و مؤلفه ۵۰۰ هرتزی که ما با تجزیۀ ال‌پی‌سی به دست آوردیم، بین سومین و چهارمین همساز قرار دارد. این نشان می‌دهد که تجزیۀ ال‌پی‌سی فرکانس قله‌های پهن طیف را پیدا می‌کند و نه صرفاً فرکانس بلندترین همسازها در این قله‌های پهن طیفی را.

تجزیۀ ال‌پی‌سی در بسیاری از نرم‌افزارهای فناوری گفتار از قبیل بازشناسی گفتار و گویشور، سنتز گفتار، رمزگذاری و انتقال گفتار به کار می‌رود. تجزیۀ ال‌پی‌سی قطعه کوتاهی از یک شکل موج آکوستیکی را انتخاب و سپس قله‌های طیف آکوستیکی، یعنی سازه‌های واکه را محاسبه می‌کند. برخلاف محاسبه زیروبمی با روش خودهمبستگی، دیرش پس‌افت‌های خودهمبستگی در ال‌پی‌سی کوتاه است (کوتاه‌تر از یک سیکل چاکنایی). از این رو، خودهمبستگی در ال‌پی‌سی به جای

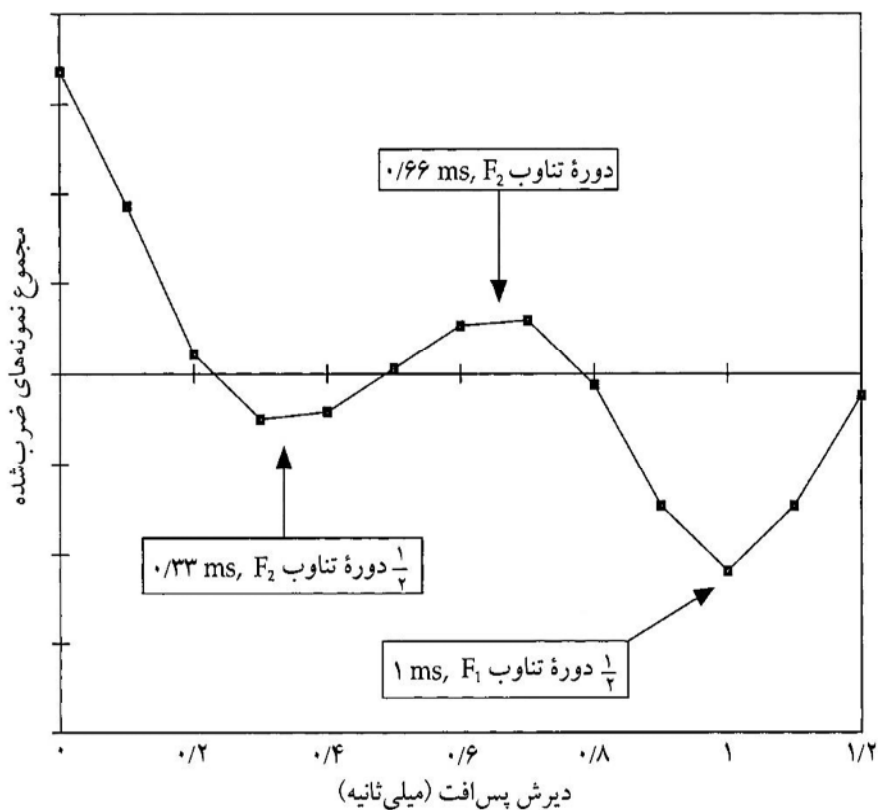
ذخیره الگوی تکرارشونده شکل موجی که واک تولید می‌کند، الگوهای تناوبی شکل موج گفتار که به وسیله بازخوان‌های سازه‌ای تولید می‌شوند، را ذخیره می‌کند.



شکل ۳-۱۶. طیف اف‌اف‌تی شکل موجی که در شکل ۳-۱۵ نشان داده شد. فلش‌ها فرکانس قله‌هایی را نشان می‌دهند که با روش خودهمبستگی در تجزیه ال‌پی‌سی بدست آمده‌اند.

ال‌پی‌سی یک «الگوریتم پیش‌گویانه خطی» نامیده می‌شود، زیرا مجموعه کوچکی از ضرایب «پیش‌بینی‌کننده» را در یک ترکیب خطی از نمونه‌های شکل موج محاسبه می‌کند. همان‌طور که خواهیم دید، این ضرایب، صافی‌ای را تعریف می‌کنند که اگر برخی مفروضات برآورده شود، در این صورت به تابع صافی دستگاه گفتار شباهت دارد. فرمول پیش‌گویانه خطی، دامنه هر نمونه اختیاری شکل موج را به صورت تابعی از چند نمونه قبل آن پیش‌بینی می‌کند. مثلاً با چهار ضریب، فرمول ال‌پی‌سی به صورت $\hat{x}_5 = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4$ است. در این فرمول a ضرایب ال‌پی‌سی و x نمونه‌های شکل موج است. بنابراین، در این مثال، مقدار نمونه $\hat{x}_5$ با محاسبه مقادیر چهار نمونه قبلی در شکل موج به دست می‌آید و ضرایب (a_i) تعیین می‌کنند که هر یک از چهار نمونه چه وزنی در فرمول دارند. توجه کنید که معادله ال‌پی‌سی شباهت بسیاری به معادله صافی‌های دیجیتال دارد که قبلاً در مورد آن بحث کردیم (معادله ۳-۲). تنها تفاوت در این است که در تجزیه ال‌پی‌سی، ضریب صافی

a_i مقدار $\hat{x}$ را از شکل موج ورودی x به جای شکل موج جدید تغییر یافته y تخمین می‌زند. هرگاه ضرایب صافی به گونه‌ای انتخاب شوند که خطای تخمین (مربع اختلاف بین $\hat{x}_5$ پیش‌بینی شده با مقدار واقعی x_5) به حداقل برسد، این ضرایب ویژگی‌های صافی دستگاه گفتار را نشان می‌دهند.

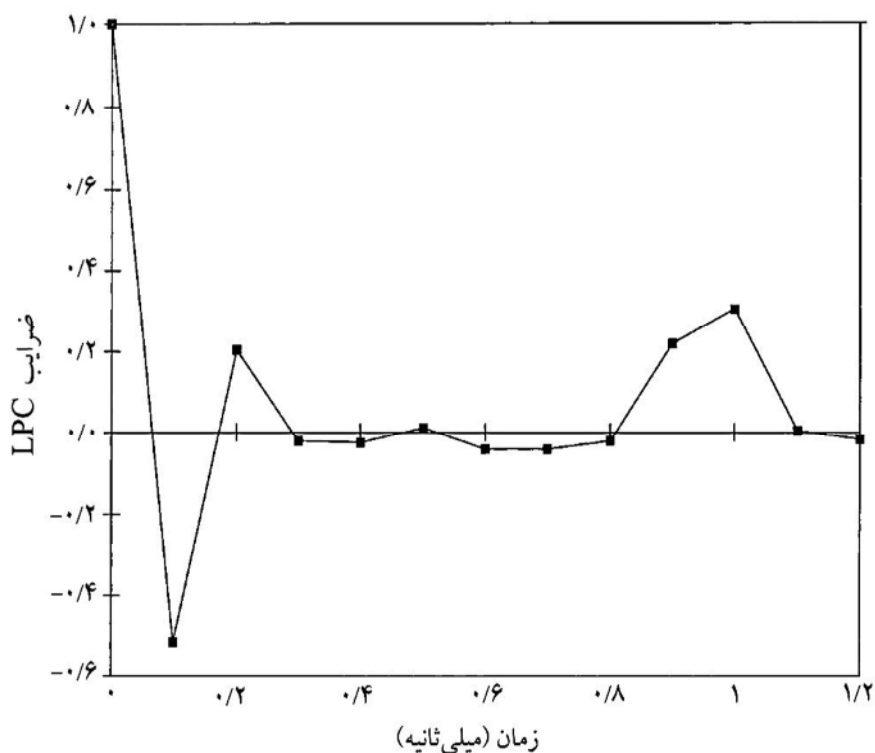


شکل ۳-۱۷. آرایه حاصل ضرب^۱ خودهمبستگی که با تجزیه الپی سی شکل موج شکل ۳-۸ به دست آمد. در این آرایه، دیرش دوره تناوب سازه‌های F_1 و F_2 به صورت قله و دره در تابع به سهولت قابل مشاهده است. اطلاعات سازه‌های بالاتر هرچند موجود اما کمتر مشخص است.

۱. product array. آرایه (array)، به طور ساده، یعنی آرایش ستونی یا ردیفی اعداد و اشیاء که در ضرب یا تقسیم به کار می‌رود. به عنوان مثال، 8×2 یک آرایه حاصل ضرب است. حاصل آرایه، در این مثال ۱۶، حاصل ضرب (product) نامیده می‌شود. (مترجم)

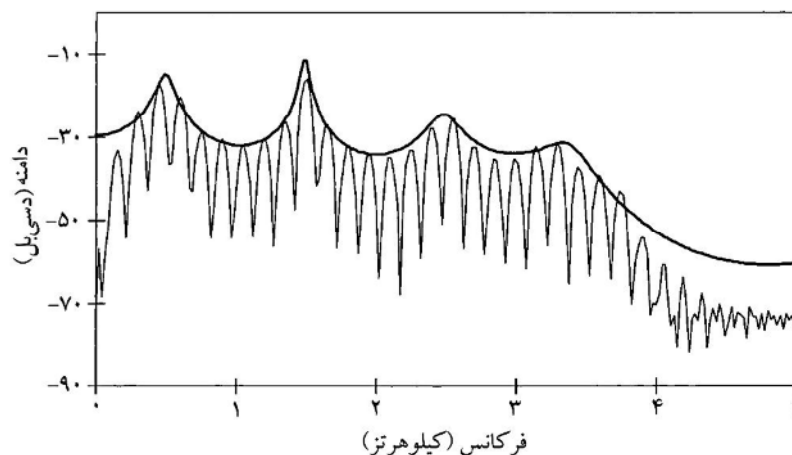
برای محاسبه ضرایب ال‌پی‌سی، نمونه‌های شکل موج از چندین سیکل زیرومی در پایان چندین پس‌افت مختلف در یکدیگر ضرب می‌شوند و فرمول آن به صورت $\sum_{n=0}^N \sum_{i=0}^M x_n x_{n-i}$ است. در این فرمول، M تعداد ضرایب، N تعداد نمونه‌ها در پنجره تجزیه و i مقدار دیرش پس‌افت است. شکل ۳-۱۷ نتایج این مرحله از الگوریتم خودهمبستگی ال‌پی‌سی (با $M=12$) را در شکل موج یک واکه نشان می‌دهد. فلش‌ها در شکل نشان می‌دهند که با ضرب نمونه‌های شکل موج در یکدیگر در فاصله پس‌افت‌های مختلف، اطلاعات دو سازه فرکانسی اول به صورت مقادیر منفی هنگامی که دیرش پس‌افت برابر با نصف دوره تناوب است و به صورت مقادیر مثبت هنگامی که دیرش پس‌افت برابر با دیرش دوره تناوب است، ذخیره می‌شود. بزرگ‌ترین مقدار وقتی به دست می‌آید که دیرش پس‌افت برابر با صفر است (یعنی هنگامی که نمونه‌های شکل موج به توان ۲ می‌رسند). هرگاه نمونه‌های شکل موج اولیه که با پس‌افت i از هم فاصله دارند با یکدیگر همبستگی داشته باشند، مقدار حاصل ضرب $x_n x_{n-i}$ مثبت است، اما اگر دیرش پس‌افت متناظر با یک همبستگی منفی باشد، مقدار حاصل ضرب منفی است.

سپس مجموعه ضرایب پیش‌بینی‌کننده a_i محاسبه می‌شوند. ضرایب حاوی اطلاعات فرکانسی هستند که در تجزیه خودهمبستگی ذخیره می‌شود. این اطلاعات به صورت یک تابع صافی خطی نشان داده می‌شوند. شکل ۳-۱۸ ضرایب ال‌پی‌سی را نشان می‌دهد که از آرایه حاصل ضرب نمایش داده شده در شکل ۳-۱۷ به دست آمده‌اند. بخش بزرگی از نمودار در یک میلی‌ثانیه است که برابر با نصف دیرش دوره تناوب F_1 اما مقادیر سایر سازه‌ها به این سهولت در مجموعه ضرایب مشخص نیست. به هر حال، همان‌طور که قبلاً نیز گفتیم، یافتن مؤلفه‌های فرکانسی یک موج مرکب تنها با بررسی دیداری شکل موج اغلب دشوار است، همان‌طور که در اینجا نیز می‌بینیم.



شکل ۳-۱۸. ضرایب ال پی سی حاصل از آرایه حاصل ضرب شکل ۳-۱۷.

اگر به طیف اف اف تی صفرگذاری شده حاصل از ضرایب ال پی سی (شکل ۳-۱۹) نگاه کنیم، به روشنی درمی یابیم که تجزیه ال پی سی اطلاعاتی درباره تابع فیلتر دستگاه گفتار را نشان می دهد. طیف ال پی سی، شکل کلی طیف بسامدی را قطع نظر از فرکانس خاص هر همساز نشان می دهد و در نتیجه، بازنمایی دقیقی از تابع فیلتر دستگاه گفتار ارائه می کند. به عبارت دیگر، قله ها در طیف ال پی سی برابرند با فرکانس های بازخوانی دستگاه گفتار (در این مثال فرکانس های بازخوانی عبارت اند از: ۵۰۰ هرتز، ۱۵۰۰ هرتز، ۲۵۰۰ هرتز، ۳۵۰۰ هرتز).



شکل ۳-۱۹. طیف‌های اف‌ا‌ف‌تی و ال‌پی‌سی واکه [ə]. شکل موج آکوستیکی این واکه در شکل ۲-۵ و طیف اف‌ا‌ف‌تی آن در شکل ۲-۶ نشان داده شد. طیف ال‌پی‌سی (خط پررنگ) که با محاسبه اف‌ا‌ف‌تی (صفرگذاری شده) از ضرایب ال‌پی‌سی نشان داده شده در شکل ۳-۱۸، تولید شد؛ قله‌های پهن طیفی واکه (قله‌سازه‌ها) را صرف نظر از فرکانس خاص هر همساز برجسته می‌کند.

ال‌پی‌سی یک ابزار بسیار مفید در آواشناسی آکوستیک محسوب می‌شود، اما برای محاسبه تابع فیلتر دستگاه گفتار محدودیت‌هایی دارد. اول اینکه، تجزیه ال‌پی‌سی فرض می‌کند که طیف واک اصولاً از قله‌های طیفی پهن، بی‌هیچ دره طیفی برجسته‌ای تشکیل یافته است (رجوع کنید به فصل ۹: پادسازه‌ها^۱). این، بدان معناست که ال‌پی‌سی برای محاسبه طیف واکه‌ها (غیرخیشومی) بسیار مناسب است، اما طیف برخی آواها مانند خیشومی‌ها، همخوان‌های کناری و برخی سایشی‌ها را که در آنها دره‌های طیفی نیز مهم هستند، به‌طور ناقصی بازنمایی می‌کند. دوم اینکه، لازم است تعداد قله‌هایی که قرار است در طیف محاسبه شوند، پیش‌تر تعیین شود. اگر تعداد قله‌های پیش‌بینی شده بیشتر از تعداد قله‌های واقعی طیف باشد، تجزیه ال‌پی‌سی قله‌های کوچکی (با پهنای باند وسیع) نشان خواهد داد که در واقع وجود ندارند و اگر تعداد قله‌های پیش‌بینی شده خیلی کم باشد، ال‌پی‌سی نخواهد توانست قله‌هایی را که واقعاً در طیف وجود دارند، ثبت کند. تعداد قله‌های طیفی که تجزیه ال‌پی‌سی در طیف بسامدی نشان می‌دهد، بر

1. anti-formants

اساس تعداد ضرایب معادله پیشگویانه خطی تعیین می‌شود. هر سازه به‌وسیله دو ضریب نشان داده می‌شود (اگر یادتان باشد در بخش ۳-۲-۱ در رابطه با فرکانس نایکویست گفتیم که دست‌کم دو نقطه برای بازنمایی حداقلی یک موج سینوسی لازم است)، به‌علاوه دو ضریب اضافی دیگر برای محاسبه شیب کلی طیف. از این رو، اگر پیش‌بینی می‌کنید که ۱۰ سازه در محدوده فرکانسی مورد نظر وجود دارد، تعداد ضرایب باید ۲۲ تا باشد $(2 + (2 \times 10))$. در عمل، تعداد سازه‌های قابل پیش‌بینی به محدوده فرکانسی مورد بررسی (که به‌نوبه خود به‌وسیله نرخ نمونه‌برداری تعیین می‌شود) و نیز به طول دستگاه گفتار گویشور بستگی دارد.

۳-۳-۶. طیف و طیف‌نگاشت

یکی از مشکلات تجزیه گفتار با استفاده از طیف قدرت، عدم بازنمایی بُعد زمان در طیف است. طیف قدرت بیشتر شبیه به یک عکس است تا یک فیلم. با استفاده از طیف قدرت می‌توان دیدگاه کاملاً دقیقی درباره مؤلفه‌های یک سیگنال در یک لحظه خاص از زمان به دست آورد، اما نمی‌توان هیچ اطلاعی درباره آنچه قبل یا بعد از آن لحظه روی داده، دریافت. مثلاً ممکن است بخواهیم ویژگی‌های پویای^۱ واژه‌های مرکب را بررسی کنیم. در این صورت، طیف‌نگاشت صوتی راه حل مسئله است. در اینجا سعی می‌کنم مفهوم آنچه را که در طیف‌نگاشت صوتی روی می‌دهد، توضیح دهم (توجه کنید که تفاوت بین اصطلاحات اسپکتروگراف و اسپکتروگرام شبیه به تفاوت بین اصطلاحات «تلگراف» و «تلگرام» است). فرض کنید که طیف واژه شوا در شکل ۳-۱۶ را تولید می‌کنید، اما به‌جای بازنمایی دامنه بر روی محور عمودی، آن را با استفاده از رنگ نشان می‌دهید. به‌عنوان مثال، در ۱۵۰ هرتز یک قله انرژی در طیف وجود دارد. فرض کنید که به‌جای رسم آن به‌صورت یک قله، خیلی ساده آن را به‌صورت یک نقطه زردرنگ (یک رنگ پرانرژی زیبا برای بیان این واقعیت که دامنه انرژی در این فرکانس بالاست) بر روی محور فرکانس

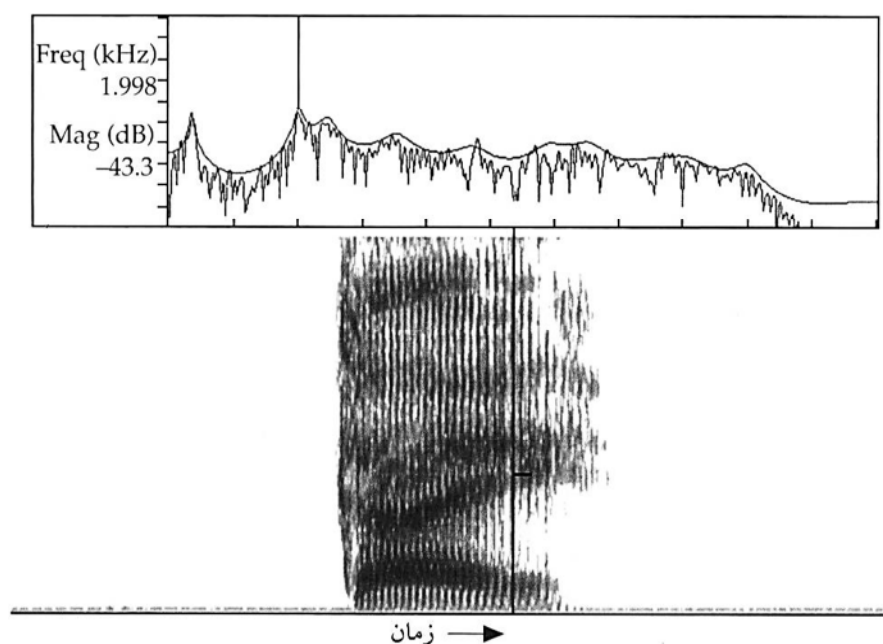
1. dynamic

علامت‌گذاری می‌کنیم. حال، فرض کنید که از یک سیستم رنگ‌بندی برای بازنمایی دامنه استفاده کنیم، به‌گونه‌ای که هر مؤلفه فرکانسی که دامنه آن در محدوده بالاترین یک‌دهم مقدار دامنه قرار گیرد، با رنگ زرد و هر مؤلفه فرکانسی که دامنه انرژی آن در محدوده یک‌دهم بعدی مقیاس دامنه قرار دارد، با رنگ قرمز و به‌همین ترتیب، مقادیر مختلف دامنه با رنگ‌بندی مشخص شوند. اکنون می‌توانیم با جایگزین کردن محور عمودی با رنگ‌های مختلف، طیفی به‌شکل یک خط رسم کنیم که برای نشان‌دادن مقادیر تقریبی دامنه مؤلفه‌ها تغییر رنگ می‌دهد. اما چرا باید زحمت استفاده از بُعدی مانند رنگ را برای بازنمایی دامنه در طیف قدرت به خود بدهیم؟ جواب این است که به این طریق می‌توانیم برای مشاهده تغییرات طیفی در طی زمان به طیف‌های بسیاری به‌طور هم‌زمان نگاه کنیم، یعنی تفاوت بین یک عکس با یک فیلم.

اکنون ببینیم چگونه می‌توان با طیف‌های رنگی یک طیف‌نگاشت ساخت. یک طیف رنگی به همان روشی که در بالا گفته شد، می‌سازیم. سپس چند میلی‌ثانیه جلوتر بر روی شکل‌موج حرکت می‌کنیم و در آن نقطه نیز یک طیف رنگی دیگر می‌سازیم. سپس این طیف را کنار طیف قبلی گذاشته و دسته‌ای از طیف‌ها را می‌سازیم. برای هم‌زمان کردن طیف‌نگاشت با موج آکوستیکی، طیف‌ها را به‌گونه‌ای تغییر جهت می‌دهیم که فرکانس بر روی محور عمودی قرار گیرد. اولین طیف در سمت چپ نمودار و هر یک از طیف‌های متوالی در سمت راست طیف قبلی خود قرار می‌گیرد. بر این اساس، در طیف‌نگاشت زمان بر روی محور افقی، فرکانس بر روی محور عمودی و دامنه به‌وسیله رنگ نشان داده می‌شود. به‌دلایل تاریخی در طیف‌نگاشت‌ها معمولاً دامنه با درجات مختلفی از رنگ خاکستری کدبندی می‌شود (بدین‌صورت که قله‌های طیف با رنگ سیاه و دره‌ها با رنگ سفید). نیمه پایین شکل ۳-۲۰ طیف‌نگاشت کلمه [ka¹] به معنای «جوجه» را نشان می‌دهد که توسط یک گویشور کانتونی^۱ تولید شده است و نیمه بالای شکل، طیف قدرت و طیف ال‌پی‌سی

۱. Cantonese. یکی از گویش‌های چینی. (مترجم)

یک نقطه از واکه مرکب [a^l] را نشان می‌دهد. خط عمودی در طیف‌نگاشت این نقطه را نشان می‌دهد. به تطابق مکان خط عمودی در طیف قدرت با مکان خط کوچک افقی در طیف‌نگاشت توجه کنید (هر دو بر روی ۱۹۹۸ کیلوهرتز قرار دارند). این، نشان می‌دهد که طیف‌نگاشت‌ها و طیف‌ها، اطلاعات فرکانسی یکسانی را نمایش می‌دهند، اما طیف‌نگاشت‌ها مثل فیلم تغییرات طیفی را در زمان نشان می‌دهند، در حالی که طیف‌ها فقط یک نقطه منفرد در زمان را نشان می‌دهند. توجه کنید که چگونه طیف‌نگاشت افزایش فرکانس دومین قله پهن طیفی، از ابتدای واکه به انتهای آن را نشان می‌دهد. این نوع تغییرات را نمی‌توان در یک طیف مشاهده نمود.



شکل ۳-۲۰. طیف‌نگاشت کلمه [ka^l] به معنای «جوجه» که توسط یک گویشور کانتونی تولید شده است. زمان بر روی محور افقی و فرکانس (از صفر تا ۵ کیلوهرتز) بر روی محور عمودی نشان داده شده است. نیمه بالایی شکل، طیف قدرت و ال‌پی‌سی یک نقطه در انتهای واکه مرکب را نشان می‌دهد (که در طیف‌نگاشت با خط عمودی مشخص شده است). خط عمودی در طیف، دومین قله پهن طیفی را نشان می‌دهد که در طیف‌نگاشت با یک خط افقی کوتاه علامت‌گذاری شده است.

معمولاً در طیف‌نگاشت‌ها تمایل داریم که همسازها را محو کنیم به گونه‌ای که بتوانیم به راحتی قله‌های پهن طیفی را مشاهده نماییم. در طیف‌نگاشت آنالوگ (پوتر و دیگران^۱ ۱۹۴۷، جوس^۲ ۱۹۴۸) این کار با تجزیه سیگنال گفتاری با استفاده از مجموعه‌ای از صافی‌های نوارگذر با پهنای باند نسبتاً وسیع صورت می‌گرفت. هر صافی فرکانس مرکزی متفاوتی داشت و به انرژی فقط در محدوده پهنای باند پاسخ می‌داد. مثلاً اگر فرکانس پایه (F_0) ۱۵۰ هرتز و پهنای باند صافی‌های طیف‌نگاشت ۳۰۰ هرتز باشد، صافی‌ها، همسازهای مجاور را در یکدیگر محو می‌کنند. از این رو، طیف‌نگاشت حاصل فقط قله‌های پهن طیف انرژی را نشان می‌دهد و نه همسازهای منفرد را.

برای تولید طیف‌نگاشت‌های دیجیتال، تجزیه افاف‌تی را برای محاسبه تک‌تک طیف‌ها به کار می‌بریم و برای کنترل پهنای باند مؤثر صافی تجزیه، تعداد نمونه‌ها را در پنجره تجزیه (با صفرگذاری) تغییر می‌دهیم. نمونه‌ای از آن در شکل ۳-۱۰ نشان داده شد. برای به دست آوردن طیف‌نگاشت‌های نوارباریک، پنجره تجزیه بلند به کار می‌بریم که در این صورت طیف‌های حاصل وضوح فرکانسی بالا و وضوح زمانی پایین دارند. اما برای به دست آوردن طیف‌نگاشت‌های نوارپهن، پنجره تجزیه کوتاه به کار می‌بریم که در این صورت، وضوح زمانی بالا (به طوری که برای مثال، اغلب می‌توانید تک‌تک پالس‌های واک را ببینید) و محوی طیفی نیز بالا خواهد بود. از بسیاری جهات، طیف‌نگاشت‌های نوارپهن در آواشناسی آکوستیک ایده‌آل محسوب می‌شوند، زیرا در عین حال که شکل کلی طیف را نشان می‌دهند، رخدادهای زمانی را به طور خیلی دقیق ثبت می‌کنند و همان‌طور که در فصل‌های بعدی خواهیم دید، ما بیشتر به مشاهده شکل کلی طیف، تا جزئیات خرد آن، علاقه‌مند هستیم.

منابعی برای مطالعه بیشتر

- Cooley, J. W., Lewis, P. A. W., and Welch, P. D. (1969) The fast Fourier transform and its applications. *IEEE Transactions on Education*, 12, 27–34.
- مقاله پژوهشی که برای اولین بار تکنیک اف‌اف‌تی را معرفی کرد.
- Joos, M. (1948) Acoustic phonetics. *Language*, 23, suppl. 1.
- یک استدلال و توجیه اولیه از اینکه چگونه تحلیل آکوستیکی (به‌ویژه طیف‌نگاشت) در جعبه ابزار آواشناسی جای گرفت.
- Lyons, R. F. (1997) *Understanding Digital Signal Processing*, Reading, MA: Addison-Wesley.
- کتاب مورد علاقه من در زمینه پردازش دیجیتال سیگنال، چون بسیار شیوا نوشته شده است.
- O'Shaughnessy, D. (1987) *Speech Communication: Human and machine*, Reading, MA: Addison-Wesley.
- یک کتاب درسی در حوزه مهندسی با پوشش گسترده مهندسی گفتار و بخش‌هایی بسیار مفید در زمینه پردازش دیجیتال سیگنال.
- Potter, R. K., Kopp, G. A., and Green, H. (1947) *Visible Speech*, Dordrecht: Van Nostrand.
- کتابی درباره چگونه «خواندن» طیف‌نگاشت، کاتالوگی بسیار مفید از آواهای گفتاری زبان انگلیسی که نشان می‌دهد وقتی این آواها کاملاً با دقت و احتیاط تولید شوند، چگونه به گوش می‌رسند.

تمرین

واژگان تخصصی

اصطلاحات زیر را تعریف کنید: پردازش دیجیتالی سیگنال (دی‌اس‌پی)، سیگنال پیوسته، سیگنال گسسته، ابزار آنالوگ، ابزار دیجیتالی، تبدیل آنالوگ به دیجیتالی، نمونه‌برداری، کمی‌سازی، نرخ نمونه‌برداری، فرکانس نایکویست، بدنمایی، صافی پادبدنمایی، بیت، نوفه کمی‌سازی، نسبت سیگنال به نوفه، برش سیگنال، پنجره شکل‌موج، خودهمبستگی، پس‌افت خودهمبستگی، نصف‌شدن زیرویمی، دوبرابرشدن زیرویمی، دامنه آرام‌اس، اف‌اف‌تی، پنجره‌گذاری، صفرگذاری، ال‌پی‌سی، پیش‌تقویت، طیف‌نگاشت.

پرسش

۱. با استفاده از یک خط‌کش، شکل‌موج نشان داده شده در شکل ۲-۳ را دیجیتالی کنید. برای این کار، فاصله‌های مساوی (با هر واحدی که مناسب است: اینچ، میلی‌متر و غیره) بر روی محور زمان و دامنه علامت‌گذاری کرده و سپس جدول زیر را کامل نمایید. فرض کنید که شکل، 0.01 میلی‌ثانیه را نشان می‌دهد، نرخ نمونه‌برداری تان چقدر است؟ تعداد گام‌ها در مقیاس کمی‌سازی تان چقدر است؟

زمان	دامنه
۰	۰

۲. فرکانس نایکویست برای هر یک از نرخ‌های نمونه‌برداری روبه‌رو چقدر است: 16000 هرتز، 11025 کیلوهرتز، 20 هرتز.

۳. فرکانس قطع یک صافی پادبدنمایی برای هر یک از نرخ‌های نمونه‌برداری روبه‌رو چقدر است: 16000 هرتز، 11025 کیلوهرتز، 44 کیلوهرتز؟

۴. اگر هر نمونه 10 بیت باشد، تعداد گام‌های دامنه چقدر خواهد بود؟ نسبت سیگنال به نوفه چقدر است؟

۵. اگر مقدار F_0 ، ۱۰۰ هرتز، ۲۰۰ هرتز و ۲۰۴ هرتز باشد، چه پس‌افتی در محاسبه زیرویمی با روش خودهمبستگی بالاترین همبستگی را (به طور ایده‌آل) ایجاد خواهد کرد؟
۶. دامنه آرام‌اس امواج سینوسی که دامنه قله آنها ۱، ۷۵، ۱۰۲۴ است، چقدر است؟
۷. دامنه آرام‌اس امواج سینوسی که دامنه قله به قله آنها ۲، ۱۰۰، ۱۸۳۴ است، چقدر است؟
۸. اگر نرخ نمونه‌برداری ۲۲ کیلوهرتز و پنجره حاوی ۵۱۲ نمونه باشد، اندازه پنجره شکل موج (برحسب میلی‌ثانیه) چقدر خواهد بود؟
۹. اگر نرخ نمونه‌برداری ۲۲ کیلوهرتز باشد، تعداد نمونه‌ها در یک پنجره ۲۰ میلی‌ثانیه‌ای چقدر است؟
۱۰. اگر نرخ نمونه‌برداری ۱۱ کیلوهرتز و طول پنجره ۵۱۲ نمونه باشد، فاصله بین نقاط طیفی در یک طیف اف‌اف‌تی چقدر خواهد بود؟
۱۱. اگر نرخ نمونه‌برداری ۲۲ کیلوهرتز و طول پنجره ۵ میلی‌ثانیه باشد، فاصله بین نقاط طیفی در یک طیف اف‌اف‌تی چقدر خواهد بود؟
۱۲. فهرست مقادیر پیش‌تقویت دامنه شکل موج دیجیتالی را که در پاسخ به سؤال اول ساختید (با استفاده از پیش‌تقویت اولین اختلاف) ارائه کنید.

فصل ۴

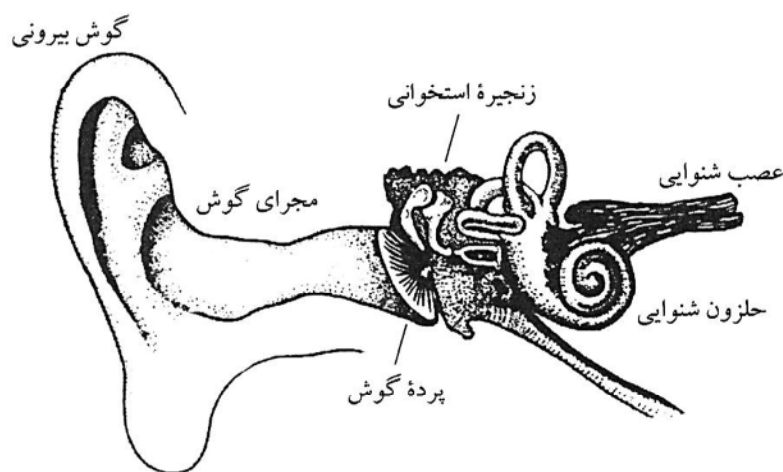
مبانی شنوایی

سیستم شنیداری انسان دستگاهی با کیفیت عالی نیست. دامنه متراکم می‌شود، فرکانس تغییر شکل^۱ می‌دهد و محو می‌شود، صداهای مجاور نیز ممکن است با هم محو شوند. از آنجا که شنوندگان عناصر شنیداری را تجربه می‌کنند و نه صورت‌های ضبط‌شده آکوستیکی مثل شکل‌موج و یا طیف‌نگاشت را، پرداختن به مبانی درک شنیداری به دلیل ارتباط آن با آکوستیک گفتار مفید خواهد بود. این فصل با مبحثی کوتاه درباره آناتومی و عملکرد سیستم شنوایی بیرونی شروع می‌شود، سپس دو تفاوت مهم بین بازنمایی آکوستیکی و شنیداری مورد بحث قرار می‌گیرد و در انتها با استفاده از شبیه‌سازی رایانه‌ای پاسخ شنیداری، تفاوت بین تجزیه آکوستیکی با تجزیه شنیداری به طور مختصر نشان داده می‌شود. در فصل‌های بعدی به مطالب مطرح‌شده در این فصل بازمی‌گردیم، زیرا در آنها درک طبقات خاصی از آواهای گفتاری مورد بحث قرار می‌گیرد.

1. warped

۴-۱. آناتومی سیستم شنوایی بیرونی

سیستم شنوایی بیرونی (آن بخش از سیستم شنوایی که خارج از مغز است) سیگنال آکوستیکی را به سیگنال عصبی تبدیل می‌کند و در طول این تبدیل، دامنه را متراکم می‌کند و نوعی تجزیه فوری بر روی سیگنال انجام می‌دهد. شکل ۴-۱ قسمت‌های اصلی سیستم شنوایی بیرونی را نشان می‌دهد (به پیکلز^۱، ۱۹۸۸ رجوع کنید). امواج صوتی به گوش بیرونی برخورد می‌کنند و از طریق مجرای گوش به سمت پرده گوش حرکت می‌کنند. پرده گوش یک پرده نازک پوستی است که در انتهای مجرای گوش مثل پوست یک طبل کشیده شده است. پرده گوش نظیر غشای میکروفن در واکنش به نوسانات فشار هوا تکان می‌خورد.



شکل ۴-۱. قسمت‌های اصلی سیستم شنوایی بیرونی. اقتباس از برودل^۲ (۱۹۶۴، شکل ۱).

این تکان‌ها سپس از طریق رشته‌ای از سه استخوان نازک در گوش میانی به گوش داخلی که پر از مایع است، هدایت می‌شود. پرده‌ای (به نام پرده بازیلا^۳) در وسط گوش داخلی که حلزونی شکل است (حلزون شنوایی) قرار دارد. این پرده در یک انتها ضخیم‌تر از انتهای دیگر است. انتهای نازک پرده بازیلا که به رشته استخوانی

1. Pickles 2. Brödel 3. basilar membrane

نزدیک‌تر است، به مؤلفه‌های فرکانس بالا در سیگنال آکوستیکی پاسخ می‌دهد، و انتهای ضخیم به مؤلفه‌های فرکانس پایین. هر رشته عصب شنوایی به بخش خاصی از پرده بازیلار می‌رود و در نتیجه اطلاعات مؤلفه فرکانسی خاصی از سیگنال آکوستیکی را انتقال می‌دهد. به این ترتیب، گوش داخلی نوعی تجزیه فوریه بر روی سیگنال آکوستیکی انجام می‌دهد و آن را به مؤلفه‌های فرکانسی مجزا تجزیه می‌کند.

۴-۲. احساس شنیداری بلندی صوت

سیستم شنوایی از طریق متراکم کردن دامنه انرژی، نوعی کنترل خودکار بر بلندی صدا اعمال می‌کند و به همین دلیل، به شدت صوت در محدوده بزرگی پاسخ می‌دهد (به مور^۱، ۱۹۸۲ رجوع کنید). به عنوان مثال، نوسانات فشاری که رعد و برق تولید می‌کند در حدود ۱۰۰۰۰۰ برابر بزرگ‌تر از نوسانات فشاری است که یک نجوا^۲ تولید می‌کند (جدول ۴-۱ را ببینید).

جدول ۴-۱. مقایسه دامنه آکوستیکی با دامنه درکی در بعضی از صداهای متداول. دامنه به سه صورت مقدار مطلق نوسان فشار هوا (میکرو پاسکال^۳)، شدت آکوستیکی (دسی بل سطح فشار صوتی^۴) و بلندی درکی (سون^۵) ارائه شده است.

تجربه شنیداری	فشار (میکرو پاسکال)	سطح دسی بل (اس پی ال)	سون
آستانه مطلق	۲۰	۰	
نجوای ضعیف	۲۰۰	۲۰	
دفتر کار آرام	۲۰۰۰	۴۰	۱
مکالمه	۲۰۰۰۰	۶۰	۴
اتوبوس شهری	۲۰۰۰۰۰	۸۰	۱۶
قطار مترو	۲۰۰۰۰۰۰	۱۰۰	۶۴
رعد بلند	۲۰۰۰۰۰۰۰	۱۲۰	۲۵۶
درد و آسیب	۲۰۰۰۰۰۰۰۰	۱۴۰	۱۰۲۴

1. Moore

3. micro-pascals (μPa)

2. Whisper

4. decibel sound pressure level (dB SPL)

5. sone

شباهت گوش داخلی به پیانو

شیوه‌ای که سیم‌های پیانو به نواخت‌ها^۱ پاسخ می‌دهد، مثالی است که منظور از عبارت «پاسخ می‌دهد» [در پاراگراف بالا] را روشن می‌کند. برای این آزمایش، یک پیانو را باز کنید تا سیم‌های آن را ببینید. برای این کار پیانوی کنسرت یا پیانوی رویال کوچک بهتر است، اما با پیانوی دیواری نیز امکان‌پذیر است. سپس پدالی که خفه‌کن‌های نمدی را از سیم‌ها جدا می‌کند، پایین بیاورید و خیلی بلند یک نُت یکنواخت بخوانید. آیا صدای ارتعاش هیچ‌یک از سیم‌ها را بعد از توقف خواندن‌تان می‌شنوید؟ اگر یک خواننده باتجربه اپرا باشید، نتیجه بهتری به دست می‌آورید. اما یک فرد مبتدی علاقه‌مند هم می‌تواند همین نتیجه را به دست آورد. از آنجا که بلندترین مؤلفه‌های سینوسی نُتی که در حال خواندن آن هستید، با فرکانس‌های بازخوانی طبیعی یک یا چند سیم پیانو مطابقت دارند، ممکن است سیم‌ها هم‌زمان با نواختی که می‌خوانید، به ارتعاش درآیند. مفهوم فرکانس بازخوانی طبیعی در مورد پردهٔ بازیلار در گوش داخلی به کار می‌رود. قسمت ضخیم پردهٔ بازیلار به‌طور طبیعی، هماهنگ با مؤلفه‌های فرکانس پایین سیگنال ورودی مرتعش می‌شود، در حالی که قسمت نازک آن به‌طور طبیعی، هماهنگ با مؤلفه‌های فرکانس بالای سیگنال به ارتعاش در می‌آید.

به مقادیر فهرست‌شده در ستون فشار جدول ۴-۱ نگاه کنید. در مورد بیشتر مردم، یک مکالمه معمولی به لحاظ شنیداری ده برابر بلندتر از یک دفتر کار آرام نیست، هرچند که بزرگی نوسانات فشار صوتی آن چنین است. به‌طور کلی، احساسات شنیداری درونی در مورد تفاوت‌های بلندی صدا برابر با اختلافات فشار صوتی نیست. مدت‌های مدیدی است که این عدم برابری بین تفاوت‌های بلندی صدا با تفاوت‌های فشار صوتی مورد توجه قرار گرفته است. به‌عنوان مثال، «استیونس» (۱۹۵۷) در یک آزمایش از شنوندگان خواست، بلندی یک صدا را طوری

1. tone

تنظیم کنند که دو برابر صدای دیگر شود، یا در آزمایش دیگری از آنها خواست بلندی صدای اول نصف صدای دوم شود. سپس پاسخ‌های شنوندگان تبدیل به مقیاسی از بلندی شنیداری شد که واحدهای آن سون نام دارد. مقیاس سون (برای شدت‌های بالای ۴۰ دسی‌بل اس‌پی‌ال) با استفاده از معادله (۴-۱) قابل محاسبه است. نمودار آن در شکل ۴-۲ (الف) رسم شده است. (دو علامت اختصاری «دسی‌بل» و «اس‌پی‌ال» در جمله بالا به کار رفت. این اصطلاحات بزودی در این بخش معرفی می‌شوند، اما برای حفظ دقت در بیان مطلب، ذکر آنها در عبارت «بالای ۴۰ دسی‌بل اس‌پی‌ال» لازم بود. «دسی‌بل» با علامت اختصاری «dB» و «سطح فشار صوتی (اس‌پی‌ال)» با علامت اختصاری «SPL» نشان داده می‌شود. در ادامه، توضیحات بیشتری در این زمینه خواهد آمد.)

$$N=2^{(dB-40)/10} \quad (4-1)$$

مقیاس سون قضاوت شنوندگان را نسبت به بلندی نسبی نشان می‌دهد. مقیاس‌گذاری برحسب سون به این صورت است که صدایی با بلندی یک دفتر کار آرام (۲۰۰۰ میکروپاسکال) ارزش ۱ دارد، صدایی که بلندی آن به لحاظ شنیداری نصف این مقدار است، ارزش ۰/۵ و صدایی که بلندی آن دو برابر است، ارزش ۲ دارد. همان‌طور که شکل نشان می‌دهد، رابطه بین فشار صوتی و بلندی صدا خطی نیست. تغییرات نسبتاً کوچک فشار صوتی در صداهای ملایم سبب تغییرات بزرگ در بلندی درکی می‌شود، در حالی که در صداهای بلند، تغییرات نسبتاً بزرگ در فشار صوتی، تغییرات کمی در بلندی درکی ایجاد می‌کند. به عنوان مثال، اگر دامنه قله از ۱۰۰۰۰۰ میکروپاسکال به ۲۰۰۰۰۰ میکروپاسکال تغییر کند، بر حسب سون، تغییر از ۱۰/۵ به ۱۶ سون خواهد بود، اما تغییر فشاری با همان بزرگی از ۲۰۰۰۰۰۰ میکروپاسکال به ۲۱۰۰۰۰۰ میکروپاسکال، بلندی صدا را کمتر از ۲ سون تغییر می‌دهد (از ۶۴ سون به ۶۵/۹ سون).

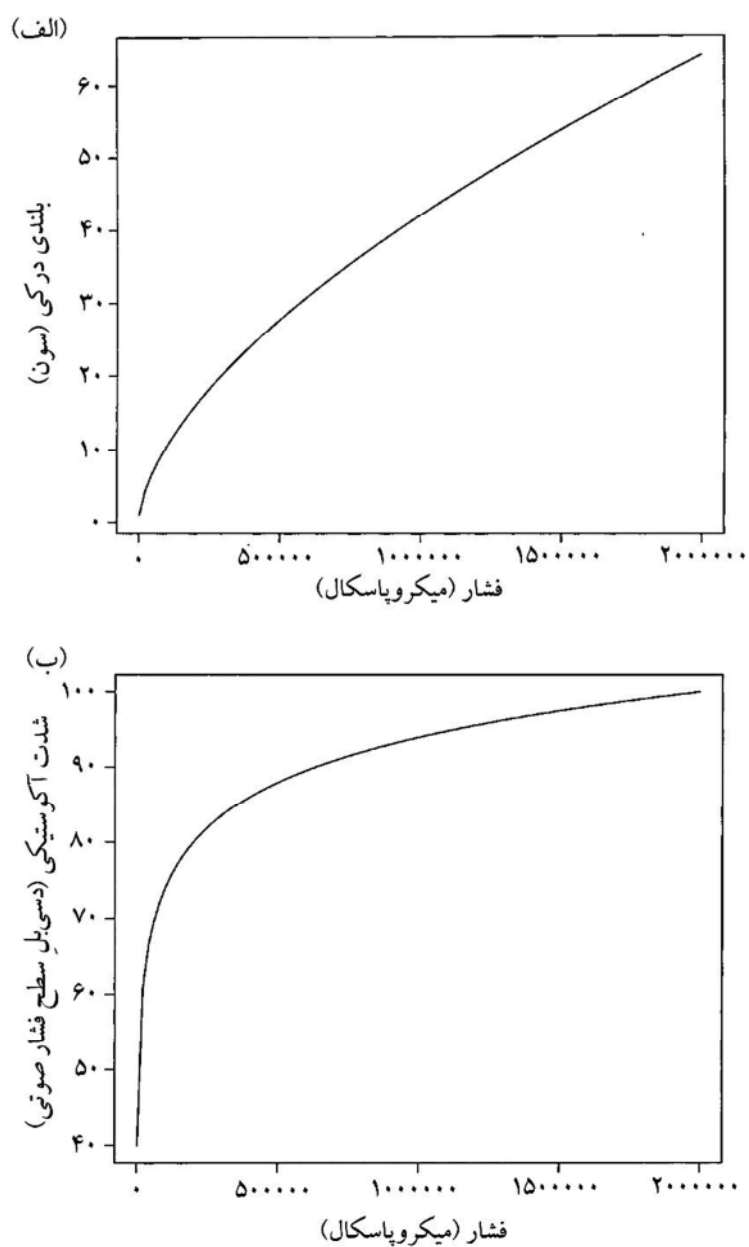
شکل ۲-۴ (ب) نیز یک مقیاس قدیمی تر را که برحسب شدت آکوستیکی نسبی است و بر اساس نام الکساندر گراهام بل نام گذاری شده، نشان می دهد. واحد بل در بسیاری از موارد خیلی بزرگ است و معمولاً به صورت یک دهم بل، یا دسی بل (به اختصار dB) به کار می رود. این مقیاس که روش محاسبه ساده ای دارد، در آواشناسی شنیداری و سایکواکوستیک^۱ به طور گسترده ای به کار می رود، زیرا تقریبی است از غیرخطی بودن احساس بلندی صدا در انسان.

همان طور که تفاوت بین دسی بل سطح فشار صوتی با دسی بل سطح احساس^۲ (به مستطیل «دسی بل» رجوع کنید) نیز نشان می دهد، بلندی درکی به عنوان تابعی از فرکانس تغییر می کند. شکل ۳-۴ رابطه بین بلندی ذهنی^۳ با دسی بل سطح فشار صوتی را نشان می دهد. منحنی این شکل، شدت صوتی دسته ای از نواخت ها را نشان می دهد که بلندی ذهنی آنها به اندازه یک نواخت ۱۰۰۰ هرتزی است که سطح فشار صوتی آن ۶۰ دسی بل است. این منحنی به تنظیمات یک همسان ساز^۴ گرافیکی در یک استریو شباهت دارد. اهرم سمت چپ همسان ساز، دامنه نسبی پایین ترین فرکانس های موسیقی را تنظیم می کند و اهرم سمت راست آن دامنه نسبی بالاترین فرکانس ها را. **منحنی بلندی همسان^۵** در شکل ۳-۴ نیز نشان می دهد که اگر می خواهید بلندی آنها به اندازه فرکانس های میانی باشد (اینکه آیا نتیجه آن خوب است یا نه، موضوع دیگری است)، لازم است شدت دامنه پایین ترین و بالاترین فرکانس ها را تقویت کنید. بنابراین، همان طور که شکل نشان می دهد، سیستم شنوایی به صداهایی که فرکانس آنها بین ۲ تا ۵ کیلوهرتز است، بیشترین حساسیت را دارد. همچنین، توجه کنید که بالای ۱۰ کیلوهرتز میزان حساسیت شنوایی خیلی سریع افت می کند. این، یکی از دلایلی است که من به نرخ نمونه برداری ۲۲ کیلوهرتز (با فرکانس نایکویست ۱۱ کیلوهرتز) برای تجزیه آوایی یا آکوستیکی توصیه می کنم.

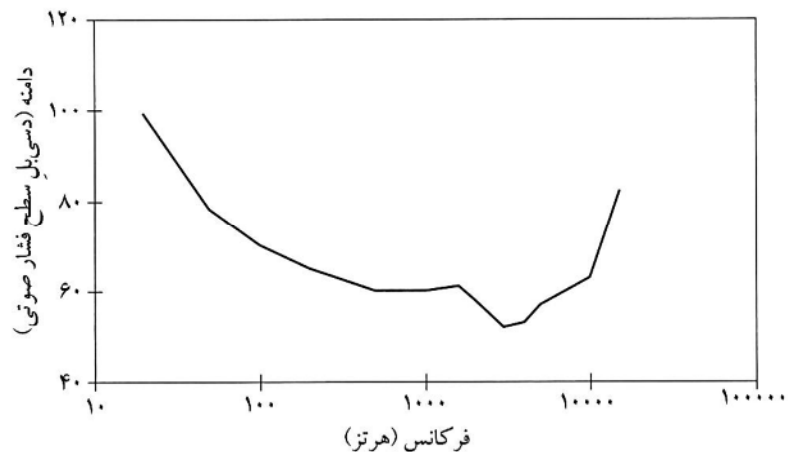
1. psycho-acoustics
3. subjective loudness

2. decibel sensation level (dB SL)
4. equalizer

5. equal loudness contour



شکل ۴-۲. بازنمایی غیرخطی بودن درک بلندی صدا با نشان دادن سطوح مختلف فشار صوتی برحسب مقیاس سون (الف) و مقیاس دسی بل (ب).



شکل ۴-۳. منحنی بلندی همسان برای نواخت‌های خالصی که با بلندگو تولید شده‌اند. منحنی، نواخت‌هایی را به هم وصل می‌کند که بلندی ذهنی آنها به اندازه یک نواخت ۱۰۰۰ هرتزی است که سطح فشار صوتی آن ۶۰ دسی بل است.

دسی بل

اگرچه معمولاً دامنه موج صوتی برحسب فشار، یا اگر انرژی آکوستیکی به انرژی الکتریکی تبدیل شود، برحسب ولت بیان می‌شود، اما مقیاس دسی بل همبستگی بهتری با بلندی درکی دارد. در این مقیاس، بلندی نسبی یک صوت برحسب شدت صوت (که نسبتی از مجذور دامنه است) با مقیاسی لگاریتمی اندازه‌گیری می‌شود. شدت آکوستیکی عبارت است از مقدار نیروی آکوستیکی که نوسان فشار موج صوتی بر واحد سطح وارد می‌کند. یک واحد متداول اندازه‌گیری شدت آکوستیکی، وات بر سانتی‌متر مربع است (W/cm^2).

یک صدا با میانگین دامنه فشار x در نظر بگیرید. از آنجا که شدت آکوستیکی نسبتی از مجذور دامنه است، بنابراین شدت x نسبت به یک صدای مرجع با دامنه فشار r برابر با x^2/r^2 است. یک بل برابر است با لگاریتم در مبنای ۱۰ این نسبت توان: $10 \log_{10}(x^2/r^2)$ ، و یک دسی بل، ۱۰ برابر این مقدار است: $10 \log_{10}(x^2/r^2)$. این فرمول در نهایت، به صورت $20 \log_{10}(x/r) = dB$ ساده می‌شود.

برای اندازه‌گیری دسی‌بل دو انتخاب رایج برای سطح مرجع r وجود دارد. یکی ۲۰ میکروپاسکال، یعنی مقدار معمول آستانه مطلق شنوایی (پایین‌ترین نوسان فشار قابل شنیدن) برای یک نواخت ۱۰۰۰ هرتزی. وقتی این مقدار مرجع به کار می‌رود، مقادیر بر حسب دسی‌بل سطح فشار صوتی بیان می‌شود. در انتخاب دیگر، برای هر فرکانس، مرجع متفاوتی برای سطح فشار به کار می‌رود. در این روش، از آستانه مطلق یک نواخت ۱۰۰۰ هرتزی به عنوان مرجع برای همه فرکانس‌ها استفاده نمی‌شود، بلکه بلندی یک نواخت نسبت به سطح آستانه مطلق معمول یک نواخت در آن فرکانس اندازه‌گیری می‌شود. وقتی این روش به کار می‌رود، مقادیر بر حسب دسی‌بل سطح احساس بیان می‌شود.

در برنامه‌های تجزیه گفتار، دامنه بر حسب دسی‌بل یا نسبت به بیشترین مقدار دامنه‌ای که یک نمونه در شکل موج دیجیتالی گفتار می‌تواند داشته باشد، بیان می‌شود که در این صورت، مقادیر دامنه اعدادی منفی هستند؛ یا نسبت به کمترین مقدار دامنه‌ای که در شکل موج دیجیتالی گفتار قابل بازنمایی است بیان می‌شود که در این صورت مقادیر دامنه، اعدادی مثبت هستند. هرگاه به دست آوردن مقدار مطلق دسی‌بل سطح فشار صوتی سیگنال ضروری نباشد، یکی از این دو امکان سطح مرجع به کار می‌رود. مثلاً وقتی هدف مقایسه بین مقادیر آرام‌اس یا دامنه طیفی نیست.

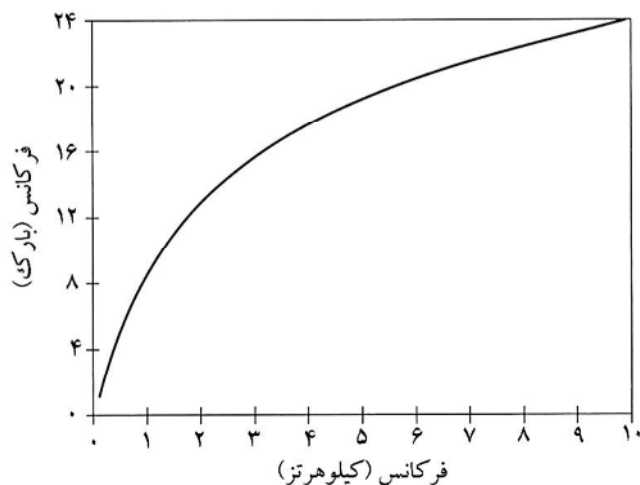
۴-۳. پاسخ فرکانسی سیستم شنوایی

همان‌طور که در بخش ۴-۱ گفته شد، سیستم شنوایی بر روی صداهای ورودی دائماً تجزیه فوریه اعمال می‌کند. اما این تجزیه فرکانسی فیزیولوژیکی شبیه روش ریاضی تجزیه فوریه سیگنال نیست. تفاوت اصلی در این است که پاسخ فرکانسی سیستم شنوایی خطی نیست. همان‌طور که یک تغییر ۱۰۰۰ میکروپاسکالی در یک صدای ملایم به لحاظ درکی معادل همین تغییر در یک صدای بلند نیست، به همین صورت، یک تغییر فرکانس از ۵۰۰ به ۱۰۰۰ هرتز نیز به لحاظ درکی معادل تغییر از

۵۰۰۰ به ۵۵۰۰ هرتز نیست. این تفاوت در شکل ۴-۴ نشان داده شده است. این شکل رابطه بین یک مقیاس فرکانس شنیداری بنام **مقیاس بارک**^۱ (زوویکر^۲، ۱۹۶۱؛ شرودر و همکاران^۳، ۱۹۷۹) با فرکانس آکوستیکی برحسب کیلوهرتز را نشان می‌دهد. «زوویکر» (۱۹۷۵) نشان داد که مقیاس بارک با یک مقیاس زیروبی درکی (بنام مقیاس مل^۴) و با فاصله در امتداد پردهٔ بازیلار تناسب دارد. فرکانس شنیداری یک نواخت ۵۰۰ هرتزی، ۴/۹ بارک است، در حالی که فرکانس شنیداری یک نواخت ۱۰۰۰ هرتزی، ۸/۵ بارک است، یعنی اختلاف ۳/۶ بارک است. از سوی دیگر، فرکانس شنیداری یک نواخت ۵۰۰۰ هرتزی، ۱۹/۲ بارک است، در حالی که فرکانس شنیداری یک نواخت ۵۵۰۰ هرتزی، ۱۹/۸ بارک است، یعنی اختلاف فقط ۰/۶ بارک است. منحنی شکل ۴-۴ این واقعیت را نشان می‌دهد که سیستم شنوایی بیشتر به تغییرات فرکانسی در انتهای پایین محدودهٔ فرکانس شنیداری حساس است تا به انتهای بالای آن.

دلیل این غیرخطی بودن احساس فرکانس این واقعیت است که تجربهٔ شنونده از زیروبی در صداهای متناوب و از طنین^۵ در صداهای مرکب تا حد زیادی به وسیلهٔ ساختار فیزیکی پردهٔ بازیلار شکل می‌گیرد. شکل ۴-۵ رابطه بین فرکانس و فاصله در طول پردهٔ بازیلار را نشان می‌دهد. همان‌طور که قبلاً ذکر شد، پایهٔ پردهٔ بازیلار نازک و رأس آن ضخیم است. در نتیجه، پایهٔ پردهٔ بازیلار به صداهای فرکانس بالا پاسخ می‌دهد و رأس آن به صداهای فرکانس پایین. همان‌طور که شکل ۴-۵ نیز نشان می‌دهد، بخش نسبتاً بزرگی از پردهٔ بازیلار به صداهای زیر ۱۰۰۰ هرتز پاسخ می‌دهد، در حالی که تنها بخش کوچکی از آن مثلاً به صداهای بین ۱۲۰۰۰ تا ۱۳۰۰۰ هرتز پاسخ می‌دهد. از این رو، تشخیص یک تغییر فرکانسی کوچک زیر ۱۰۰۰ هرتز آسان‌تر است تا تشخیص یک تغییر کوچک فرکانسی بالای ۱۲۰۰۰ هرتز. رابطه بین فرکانس شنیداری و فرکانس آکوستیکی که در شکل ۴-۴ نشان داده شده، ناشی از ساختار پردهٔ بازیلار در گوش داخلی است.

1. Bark scale 2. Zwicker 3. Schroeder 4. Mel scale 5. timbre

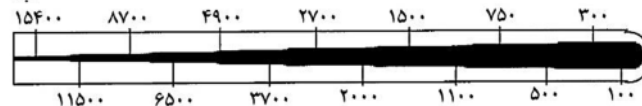


شکل ۴-۴. مقایسه یک مقیاس فرکانس شنیداری (مقیاس بارک) با یک مقیاس فرکانس آکوستیکی (به کیلوهرتز) نشان می‌دهد که سیستم شنوایی به تغییرات فرکانسی کوچک در انتهای فرکانس پایین محدوده شنیداری حساس‌تر است.

(الف)



(ب)



شکل ۴-۵. (الف) گوش داخلی. (ب) محل تقریبی حساسیت به برخی فرکانس‌ها در امتداد پردهٔ بازیلا در حالتی که گوش داخلی از حالت حلقه در آمده است. برای مشاهدهٔ موقعیت گوش داخلی در ارتباط با سایر بخش‌های سیستم شنوایی بیرونی به شکل ۴-۱ رجوع کنید.

۴-۴. اشباع^۱ و استتار^۲

نورون‌های حسی گوش داخلی و عصب شنوایی که به صدا واکنش نشان می‌دهند، ماشین‌های فیزیولوژیکی هستند که با ذخایر شیمیایی (یون‌های باردار سدیم و پتاسیم) کار می‌کنند. لذا، وقتی ذخیره آنها کم می‌شود یا با حداکثر ظرفیت کار می‌کنند، ممکن است نتوانند با تمام توان به صدا واکنش نشان دهند. در واقع، این وضعیت بسیار پیش می‌آید.

به عنوان مثال، پس از مدت زمان کوتاهی سکوت، واکنش سلول عصب شنوایی به یک نواخت، خیلی بزرگ‌تر است نسبت به وقتی که از اجرای نواخت مدت کوتاهی می‌گذرد. طی سکوت، نورون‌ها با دریافت یون‌های باردار مثبت دوباره به طور کامل «باتری‌شان شارژ می‌شود». بنابراین، وقتی یک صدا، پردهٔ بازیلار در حلزون گوش را تکان می‌دهد، سلول‌های مویی و نورون‌های عصب شنوایی آمادهٔ شلیک [ارسال پیام الکتریکی] هستند. طول این مدت زمان حساسیت شدید نورون‌ها، از نورونی به نورون دیگر فرق می‌کند، اما معمولاً کوتاه است، شاید ۵ تا ۱۰ میلی‌ثانیه. جالب این است که طول آن تقریباً به اندازهٔ دیرش انفجار رهش^۳ یک همخوان انسدادی است و به نظر می‌رسد که حساسیت بیشتر نورون‌های شنوایی پس از مدت زمان کوتاهی سکوت، سبب افزایش حساسیت درکی فرد نسبت به اطلاعات انفجار رهش انسدادی می‌شود. همین مکانیسم احتمالاً باید امکان شنیدن ویژگی‌های متغیر صوتی را به صورت کلی‌تری میسر سازد، زیرا سکوت (تا آنجا که به نورون‌ها مربوط می‌شود) یعنی فقدان انرژی آکوستیکی در فرکانس مرکزی خاصی از سلول عصب شنوایی. بنابراین، یک فعالیت شدید اولیه سبب کاهش میزان اطلاعی بودن صداها^۴ ایستا می‌شود، چه سکوت باشد چه نباشد.

جنبهٔ دیگری از اینکه صدا چگونه در زمان به وسیلهٔ سیستم شنوایی ثبت می‌شود، پدیدهٔ استتار است. در استتار حضور یک صدا سبب می‌شود صدای مجاور

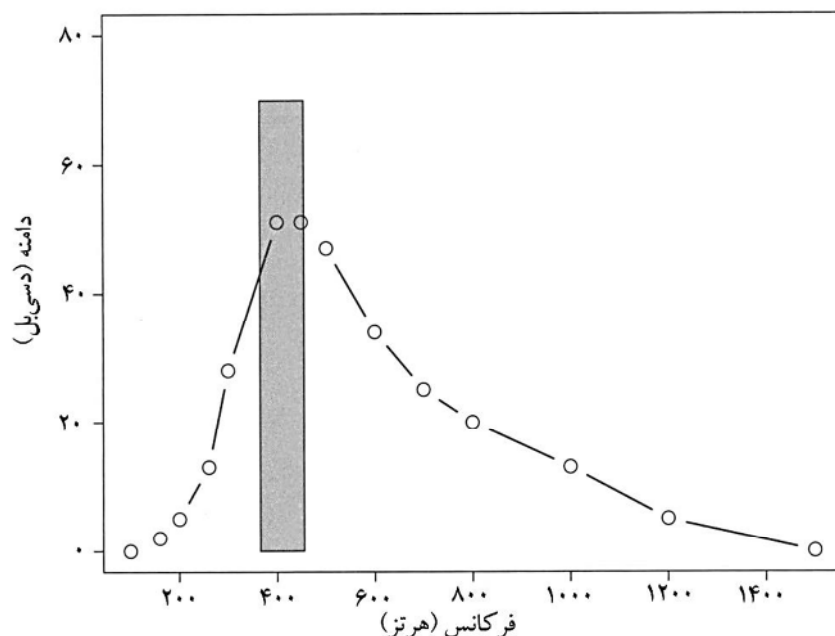
1. saturation

2. masking

3. release burst

4. steady-state

دیگر به سختی شنیده شود. استتار یک اثر «شلوغی خط»^۱ نامیده می‌شود. منظور این است که اگر یک نورون در واکنش به یک صدا شلیک کند [یک پیام الکتریکی بفرستد] و ارسال صدای دیگری نیز بخواهد با شلیک همان نورون صورت بگیرد، در این صورت صدای دوم نمی‌تواند شلیک قابل توجهی ایجاد کند. در نتیجه، سیستم به صدای دوم نسبتاً غیرحساس خواهد بود. در ادامه دو نوع استتار که مربوط به درک گفتار هستند، مورد بحث قرار خواهد گرفت: «استتار فرکانسی» و «استتار زمانی».



شکل ۴-۶. بازنمایی استتار هم‌زمان که گسترش روبه‌بالای استتار را نشان می‌دهد. نوار خاکستری، فرکانس و دامنه یک نوفه استتارکننده را نشان می‌دهد و نقطه‌ها، بالارفتن آستانه شنیدن نواخت‌های آزمایش در فرکانس‌های مختلف را نشان می‌دهد. این نواخت‌ها هم‌زمان با نوفه استتارکننده تولید شده‌اند. شکل، داده‌های اگان و هیک^۲ (۱۹۵۰) را به شکل دیگری نشان می‌دهد.

شکل ۴-۶ استتار فرکانسی را نشان می‌دهد. شکل، نتیجه آزمایشی که با یک باند کوتاه نوفه (نوار خاکستری در شکل) و مجموعه‌ای از نواخت‌های سینوسی

1. line busy

2. Egan and Hake

(دایره‌های توخالی) انجام شد، را نشان می‌دهد. در این شکل، پهنای نوفه استتارکننده ۹۰ هرتز، فرکانس مرکزی آن ۴۱۰ هرتز و بلندی آن ۷۰ دسی‌بل سطح فشار صوتی است. نقطه‌ها نشان می‌دهد که چقدر دامنه یک نواخت باید افزایش پیدا کند تا با وجود نوفه استتارکننده قابل شنیدن باشد (به عبارت دیگر، نقاط، بالارفتن سطح آستانه شنیدن نواخت‌های آزمایش را نشان می‌دهند). به عنوان مثال، به هیچ وجه آستانه دامنه یک نواخت ۱۰۰ هرتزی (اولین نقطه)، تحت تأثیر حضور استتارکننده ۴۱۰ هرتزی قرار نگرفته است، اما دامنه یک نواخت ۴۰۰ هرتزی باید تا ۵۰ دسی‌بل تقویت شود تا قابل شنیدن باشد. یک جنبه اصلی در استتار فرکانسی شکل ۴-۶، گسترش روبه‌بالای استتار^۱ نام دارد. یعنی نواخت‌ها در فرکانس‌های بالاتر از نوفه استتاری بیشتر تحت تأثیر قرار گرفته‌اند تا نواخت‌های زیر نوفه استتاری. بنابراین، برای شنیدن یک نواخت ۶۱۰ هرتزی (۲۰۰ هرتز بالاتر از فرکانس مرکزی نوفه استتاری) نواخت باید تا ۳۸ دسی‌بل بیشتر از آستانه معمول بلندی خود تقویت شود، در حالی که در یک نواخت ۲۱۰ هرتزی (۲۰۰ هرتز پایین‌تر از فرکانس مرکزی نوفه استتاری) اصولاً^۲ به ندرت به تقویتی نیاز است. این نشان می‌دهد که نوفه‌های فرکانس پایین گرایش به استتار مؤلفه‌های فرکانس بالا را دارند.

گسترش روبه‌بالای استتار: به چه دلیل و برای چه؟

دو نکته درباره گسترش روبه‌بالای استتار قابل گفتن است. اول اینکه، دلیل آن احتمالاً مکانیک ارتعاش پرده بازilar در حلزون شنوایی است. موج فشار یک موج سینوسی که از استخوان‌های گوش میانی به حلزون شنوایی منتقل شده در طول پرده بازilar حرکت می‌کند، دامنه آن افزایش پیدا می‌کند (یعنی پرده بازilar را بیشتر و بیشتر تکان می‌دهد) تا رسیدن به نقطه حداکثر پاسخ به فرکانس موج سینوسی (شکل ۴-۵ را ببینید). سپس ارتعاش پرده بازilar به سرعت متوقف می‌شود. این مکانیسم سبب می‌شود

←

1. upward spread of masking

پایه پردهٔ بازیلار که فرکانس‌های بالا را ثبت می‌کند، به وسیلهٔ صداهایی با فرکانس پایین تحریک شود، در حالی که رأس پردهٔ بازیلار که فرکانس‌های پایین را ثبت می‌کند، چندان تحت تأثیر صداهای فرکانس بالا قرار نگیرد. بنابراین، گسترش روبه بالای استتار، محصول فیزیوژیکی عملکرد مکانیکی این حلقهٔ مملو از مایع در سر است.

ممکن است بگویید «برای چه؟». گسترش روبه بالای استتار برای فشرده سازی صوت در امپی‌تری‌ها به کار می‌رود. در فصل ۳ به فشرده سازی صوت اشاره کردیم و گفتیم که داده‌های صوتی خام، پهنای باند زیادی اشغال می‌کنند. در استاندارد فشرده سازی امپی‌تری برای حذف گزینشی مؤلفه‌های فرکانسی از داده‌های فشرده شده، از روش استتار و به ویژه از گسترش روبه بالای استتار استفاده می‌شود. آن بیت‌هایی را که در فرکانس‌های بالای شکل ۴-۶ در هر صورت قادر به شنیدن آنها نیستید، حذف کنید! و فضای حافظه را خیلی ساده با حذف بیت‌های غیرقابل شنیدن حفظ نمایید.

دومین نوع استتار، استتار زمانی است. منظور از استتار زمانی این است که صداهایی که به دنبال هم می‌آیند ممکن است یکدیگر را مبهم و نامشخص کنند. به عنوان مثال، یک صدای ملایم کوتاه ممکن است به خودی خود کاملاً قابل شنیدن باشد، اما اگر بلافاصله به دنبال یک صدای خیلی بلندتر در همان فرکانس تولید شود، ممکن است کاملاً مبهم و نامشخص شود. این پدیدهٔ «استتار پیشرو» مستلزم پارامترهای بسیاری است. دامنهٔ نوفهٔ استتارکننده برای آنکه بر درک گفتار تأثیر بگذارد، باید بیش از ۴۰ دسی‌بل سطح فشار صوتی داشته باشد و فرکانس صدای استتار شده نیز باید هماهنگ با فرکانس صدای استتارکننده (یا در زیرمجموعهٔ فرکانسی آن) باشد. اثر استتار خیلی سریع پایین می‌آید و بعد از حدوداً ۲۵ میلی ثانیه تقریباً هیچ تأثیر عملی ندارد. در گفتار، تأثیر جزئی استتار پیشرو را می‌توان در پایانه‌های^۱ واکه دید (واکه‌ها بلندترین صداهای گفتاری هستند). از سویی استتار پسرو، که طی آن قابلیت شنیدن یک صدای ملایم به واسطهٔ تولید یک صدای بلند

1. offset

پس از آن کاهش می‌یابد، حتی ارتباط زیادی با درک گفتار ندارد، اگرچه خود معمای جالبی است که چگونه چیزی می‌تواند در زمان به عقب بازگردد و بر درک ما تأثیر بگذارد. در واقع جادو نیست، تنها دلیل آن این است که سیگنال‌های قوی خیلی سریع‌تر از سیگنال‌های ضعیف در سیستم عصبی حرکت می‌کنند.

۴-۵. بازنمایی‌های شنیداری

در واقع، منظور از آنچه گفتیم این است که وقتی طیف قدرت یک آوای گفتاری را محاسبه می‌کنیم، مقیاس‌های فرکانس و بلندی صوت در ابزار مورد استفاده (مثلاً کامپیوتر یا طیف‌نگار) همانند مقیاس‌های فرکانس و بلندی در سیستم شنوایی نیست. از این رو، تجزیه آکوستیکی آواهای گفتاری ممکن است با تجربه شنونده مطابقت نداشته باشد. این عدم تطابق به‌ویژه در صداهایی مثل انفجار رهش بعضی از انسدادی‌ها و سایشی‌ها که مقادیر زیادی انرژی در فرکانس‌های بالا دارند و یا با تغییرات ناگهانی در دامنه همراه هستند، شگرف است. یک راه برای جلوگیری از این عدم تطابق بین تجزیه آکوستیکی با تجربه شنونده، پیاده‌سازی یک مدل کاربردی از سیستم شنوایی است. بعضی از موارد استفاده از مدل‌های شنیداری در تجزیه گفتار عبارت‌اند از لیلجنکرائنتس و لیندبلوم^۱ (۱۹۷۲) بلادن و لیندبلوم^۲ (۱۹۸۱)، جانسون^۳ (۱۹۸۹)، لاینز^۴ (۱۹۸۲)، پترسون^۵ (۱۹۷۶)، مور و گلاسبرگ^۶ (۱۹۸۳)، و سنف^۷ (۱۹۸۸). شکل ۴-۷ تفاوت بین طیف شنیداری با طیف آکوستیکی یک موج مرکب متشکل از مؤلفه‌های سینوسی ۵۰۰ هرتزی و ۱۵۰۰ هرتزی را نشان می‌دهد. محور عمودی دامنه انرژی برحسب دسی‌بل را نشان می‌دهد و محور افقی، در قسمت پایین نمودار، فرکانس برحسب هرتز و در بالای نمودار، فرکانس برحسب بارک را نشان می‌دهد. من این طیف شنیداری و سایر طیف‌هایی را که در شکل‌های بعدی می‌بینید، با یک برنامه کامپیوتری (جانسون، ۱۹۸۹) که ویژگی‌های پاسخ فرکانسی شکل ۴-۴ و منحنی بلندی همسان شکل ۴-۳ را تقلید می‌کند، رسم کردم. دقت کنید از آنجا که

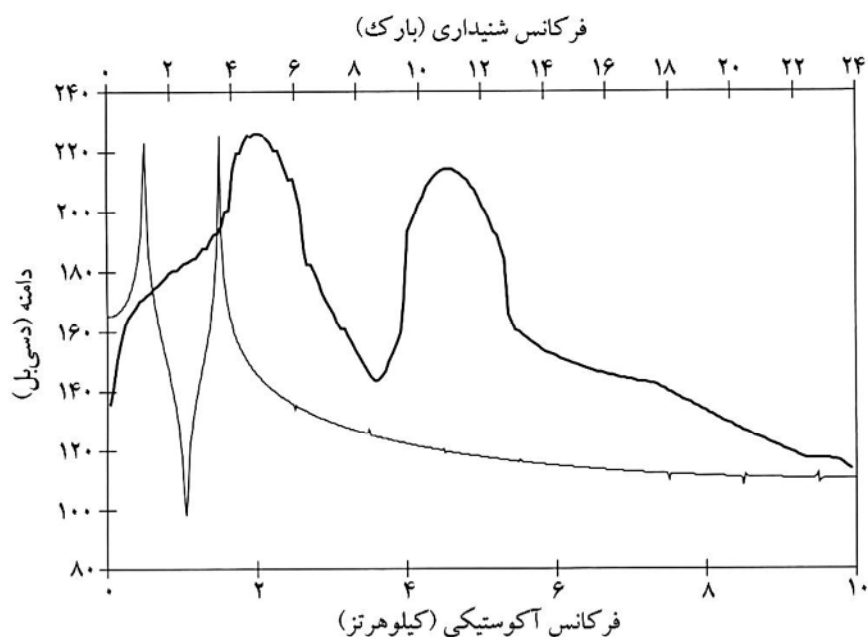
1. Liljencrants and Lindblom
5. Patterson

2. Bladon and Lindblom
6. Moore and Glasberg

3. Johnson
7. Seneff

4. Lyons

مقیاس فرکانس آکوستیکی با مقیاس فرکانس شنیداری متفاوت است، قله‌ها در این دو بازنمایی در جاهای متفاوتی قرار گرفته‌اند، هرچند که هر دو طیف، محدوده فرکانسی از صفر تا ۱۰۰۰۰ هرتز را پوشش می‌دهند. تقریباً نیمی از مقیاس فرکانس شنیداری، فرکانس‌های زیر ۱۵۰۰ هرتز را پوشش می‌دهد، در حالی که همین محدوده، کمتر از دو دهم بازنمایی آکوستیکی را در بر می‌گیرد. بنابراین، مؤلفه‌های دارای فرکانس پایین بر طیف شنیداری حاکم هستند. همچنین، توجه کنید که در طیف شنیداری وجود نوعی محوی فرکانسی^۱ سبب می‌شود قله ۱۱ بارکی (۱۵۰۰ هرتزی) تاحدی پهن‌تر از قله ۵ بارکی (۵۰۰ هرتزی) باشد. این اثر محوی طیفی با افزایش فرکانس افزایش پیدا می‌کند.



شکل ۴-۷. مقایسه طیف آکوستیکی (نمودار کم‌رنگ) با طیف شنیداری (نمودار پررنگ) یک موج مرکب که از موج‌های سینوسی ۵۰۰ و ۱۵۰۰ هرتزی تشکیل یافته است. هر دو طیف از صفر تا ۱۰ کیلوهرتز امتداد دارند، البته روی مقیاس‌های فرکانسی متفاوت. طیف شنیداری از طیف آکوستیکی با استفاده از مدل جانسون (۱۹۸۹) محاسبه شد.

1. frequency smearing

شکل ۴-۸ نمونه‌ای از تفاوت بین طیف آکوستیکی با طیف شنیداری گفتار را نشان می‌دهد. نمودار (الف) طیف‌های آکوستیکی رهش بست در سه نچ‌آوای زبان خوسا^۱ و نمودار (ب) طیف‌های شنیداری این آواها را نشان می‌دهد. این شکل نیز مانند شکل ۴-۷ نشان‌دهنده تفاوت‌های متعدد بین طیف آکوستیکی با طیف شنیداری است. اول اینکه، ناحیه بین ۶ تا ۱۰ کیلوهرتز (۲۴-۲۰ بارک در طیف شنیداری) که در آن تفاوت بین نچ‌آواها چندان نیست، در طیف‌های شنیداری نیز چندان برجسته نیست. این بخش بی‌معنا در طیف‌های آکوستیکی دوپنجم مقیاس فرکانس را در بر می‌گیرد، در حالی که فقط یک‌پنجم مقیاس فرکانس شنیداری را شامل می‌شود. این سبب می‌شود که تفاوت بین طیف‌ها به لحاظ بصری و احتمالاً شنیداری افزایش یابد. دوم اینکه، طیف‌های شنیداری قله‌های خیلی کمتری نسبت به طیف‌های آکوستیکی دارند. در این رابطه لازم به ذکر است که برای هموارکردن طیف‌های آکوستیکی شکل ۴-۸، محاسبه آنها با استفاده از تجزیه ال‌پی‌سی صورت گرفت. طیف‌های افافتی که برای محاسبه طیف‌های شنیداری به کار رفت، خیلی پیچیده‌تر از این طیف‌های هموار ال‌پی‌سی هستند. البته همواری آشکار طیف‌های شنیداری نیز به دلیل افزایش پهنای باند صافی‌های شنیداری در فرکانس‌های بالاست.

مدل‌های شنیداری از این نظر که راهی برای نگاه به سیگنال‌های گفتاری از منظر شنونده هستند، جالب‌اند. در آواشناسی میزان مفیدبودن مدل‌های شنیداری بستگی به میزان دقت در شبیه‌سازی سیستم شنوایی بیرونی دارد. از این رو، بازنمایی‌های این کتاب به وسیله مدل‌های شنیداری که روابط غیرخطی معتبر و بسیار مطالعه‌شده بلندی شنیداری و پاسخ فرکانسی را پیاده می‌کنند، تولید شد و از حوزه‌هایی از دانش که خیلی کمتر به سیگنال‌های پیچیده‌ای مانند گفتار پرداخته‌اند، خودداری شد.

این بازنمایی‌های نسبتاً محتاطانه شنیداری حاکی از آن هستند که تجزیه‌های آکوستیکی تنها تخمینی اجمالی از بازنمایی‌های شنیداری که شنوندگان برای تشخیص آواهای گفتاری به کار می‌برند، به دست می‌دهند.

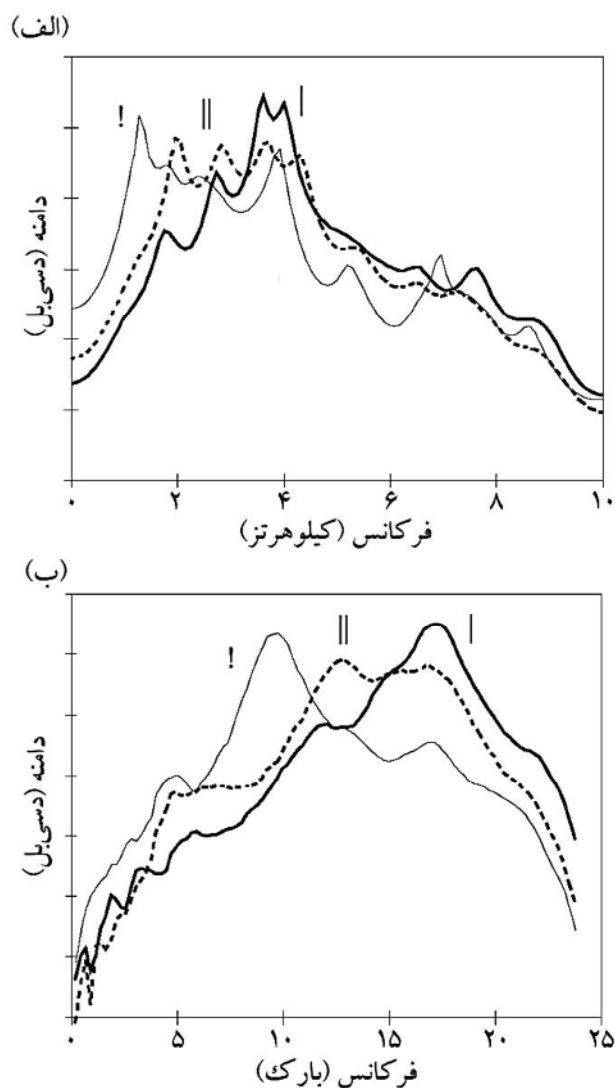
۱. Xhosa. یکی از زبان‌های رسمی آفریقای جنوبی. (مترجم)

در فصل سوم گفتیم که طیف‌نگاشت‌های دیجیتالی با ذخیره دامنۀ طیفی حاصل از مجموعه‌ای از طیف‌های اف‌اف‌تی به صورت سایه‌های خاکستری به دست می‌آیند. همین روش بازنمایی می‌تواند برای ایجاد طیف‌نگاشت‌های شنیداری از رشته‌ای از طیف‌های شنیداری نیز به کار رود. شکل ۴-۹ طیف‌نگاشت آکوستیکی و طیف‌نگاشت شنیداری واژه کانتونی [ka'] «جوجه» را نشان می‌دهد (شکل ۳-۲۰ را ببینید). برای به دست آوردن این شکل، من از یک مدل شنیداری قابل دسترس در اینترنت استفاده کردم (مدل حلزونی^۱ لاینز (لاینز، ۱۹۸۲؛ اسلنی^۲، ۱۹۸۸) که در سایت <http://linguistics.berkeley.edu/phonlab/resources/> در دسترس است). طیف‌نگاشت شنیداری که حلزون‌نگاشت^۳ نیز نامیده می‌شود، ویژگی‌های طیف‌ها و طیف‌نگاشت‌های شنیداری را ترکیب می‌کند. در حلزون‌نگاشت نیز مثل طیف‌نگاشت، پاسخ شنیداری با رسم دامنۀ طیفی به صورت سایه‌های خاکستری بازنمایی شده است. بدین صورت که زمان بر روی محور افقی و فرکانس بر روی محور عمودی است. توجه کنید اگرچه هر دو بازنمایی محدوده فرکانسی یکسانی را پوشش می‌دهند (صفر تا ۱۱ کیلوهرتز)، اما در حلزون‌نگاشت به دلیل مقیاس فرکانسی شنیداری، حرکات پایین‌ترین نقاط تمرکز انرژی طیفی واکه به لحاظ بصری بسیار برجسته‌ترند.

1. cochlear model

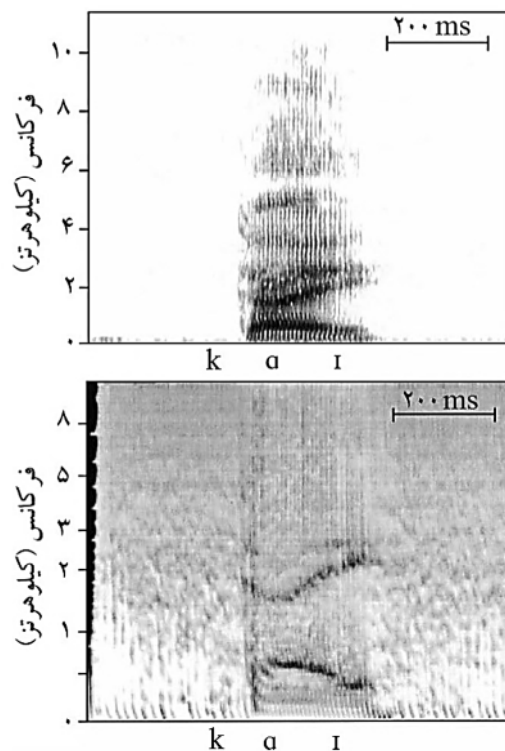
2. Slaney

3. cochleagram



شکل ۴-۸. مقایسه طیف قدرت آکوستیکی (الف) با طیف شنیداری (ب) انفجار رهش در سه نهج آوای زبان «خوسا». علامت این آواها در آی پی ای^۱ به صورت [!] دندانی، [!] پس لثوی، و [||] کناری است. در طیف های آکوستیکی، تفاوت نهج آواها در فرکانس های کمتر از ۶ کیلوهرتز است. در طیف های شنیداری، فرکانس های غیر تمایز دهنده بالای ۶ کیلوهرتز در انتهای مقیاس، متراکم شده اند (بالای ۲۰ بارک). روی محور عمودی فاصله بین علامت گذاری ها ۱۰ دسی بل است.

1. International Phonetic Association (IPA)



شکل ۴-۹. مقایسه طیف‌نگاشت آکوستیکی (بالا) با طیف‌نگاشت شنیداری یا حلزون‌نگاشت (پایین) در واژه کانتونی [ka'] «جوجه». حلزون‌نگاشت با مدل حلزونی لاینز (۱۹۸۲) تولید شده است.

منابعی برای مطالعه بیشتر

- Bladon, A. and Lindblom, B. (1981) Modeling the judgement of vowel quality differences. *Journal of the Acoustical Society of America*, 69, 1414-22.
 استفاده از یک مدل شنیداری برای پیش‌بینی نتایج درک واژه.
- Brödel, M. (1946) *Three Unpublished Drawings of the Anatomy of the Human Ear*, Philadelphia: Saunders.
- منبع شکل ۴-۱.
- Johnson, K. (1989) Contrast and normalization in vowel perception. *Journal of Phonetics*, 18, 229-54.
 استفاده از یک مدل شنیداری برای پیش‌بینی نتایج درک واژه.
- Liljencrants, J. and Lindblom, Björn (1972) Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48, 839-62.

- استفاده از یک مدل شنیداری برای پیش‌بینی الگوهای زبان‌گذر در فهرست‌های واکه‌ای.
- Lyons, R. F. (1982) A computational model of filtering, detection and compression in the cochlea. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 1282-5.
- شبیه‌سازی تبدیل سیگنال‌های آکوستیکی به سیگنال‌های عصبی شنیداری. در این کتاب از مدل اسلنی (۱۹۸۸) برای تولید «حلزون‌نگاشت‌ها» استفاده شده است.
- Moore, B. C. J. (1982) *An Introduction to the Psychology of Hearing*, 2nd edn., New York: Academic Press.
- یک کتاب مقدماتی جامع راجع به سنجش رفتاری توانایی شنیداری انسان - سایکوفیزیک شنوایی.
- Moore, B. C. J. and Glasberg, B. R. (1983) Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America*, 74, 750-3.
- توصیف یک مقیاس فرکانسی شنیداری به نام پهنای باند مستطیلی هم‌ارز^۱ (ای آر بی) و توضیح اینکه چگونه می‌توان طیف‌های شنیداری مصنوعی را با این اطلاعات محاسبه کرد.
- Patterson, R. D. (1976) Auditory filter shapes derived from noise stimuli, *Journal of the Acoustical Society of America*, 59, 640-54.
- نشان می‌دهد چگونه می‌توان با استفاده از یک بانک فیلتر شنیداری، طیف‌های شنیداری مصنوعی را محاسبه نمود.
- Pickles, J. O. (1988) *An Introduction to the Physiology of Hearing*, 2nd edit, New York: Academic Press.
- یک کتاب مقدماتی معتبر و جالب در زمینه فیزیک و شیمی شنوایی به علاوه مروری مفید بر سایکوفیزیک شنوایی.
- Schroeder, M. R., Atal, B. S., and Hall, J. L. (1979) Objective measure of certain speech signal degradations based on masking properties of human auditory perception. In B. Lindblom and S. Ohman (eds.), *Frontiers of Speech Communication Research*, London: Academic Press, 217-29.
- اندازه‌گیری و کاربرد مقیاس فرکانسی بارک.
- Seneff, S. (1988) A joint synchrony/mean-rate model of auditory speech processing. *Journal of Phonetics*, 16, 55-76.
- Slaney, M. (1988) Lyons' cochlear model. *Apple Technical Report*, 13. Apple Corporate Library, Cupertino, CA.
- مدلی رایج از شبیه‌سازی شنوایی ریچارد لاینز (۱۹۸۲).
- Stevens, S. S. (1957) Concerning the form of the loudness function. *Journal of the Acoustical Society of America*, 29, 603-6.
- مثالی از سایکوفیزیک شنوایی کلاسیک دهه ۱۹۵۰ میلادی که در آن مقیاس سون برای اولین بار توصیف شد.
- Zwicker, E. (1961) Subdivision of the audible frequency range into critical bands (*Frequenzgruppen*). *Journal of the Acoustical Society of America*, 33, 248.
- یک توصیف اولیه از مقیاس بارک باندهای فرکانسی بحرانی شنیداری.
- Zwicker, E. (1975) Scaling. In W. D. Keidel and W. D. Neff (eds.), *Auditory System: Physiology (CNS), behavioral studies, psychoacoustics*, Berlin: Springer-Verlag.
- مروری اجمالی بر مقیاس‌های متعدد بلندی و فرکانس که در پژوهش‌های سایکوفیزیک شنوایی به کار می‌رود.

1. equivalent rectangular bandwidth (ERB)

تمرین

واژگان تخصصی

اصطلاحات زیر را تعریف کنید: سیستم شنوایی بیرونی، پرده گوش، گوش میانی، گوش داخلی، پرده بازیلار، حلزون گوش، فرکانس بازخوانی طبیعی، سون، دسی بل، منحنی بلندی همسان، دسی بل سطح فشار صوتی، دسی بل سطح احساس، مقیاس بارک، گسترش روبه بالای استتار، مدل شنیداری، طیف شنیداری، حلزوننگاشت.

پرسش

۱. با توجه به شکل ۴-۲، سطح فشار صوتی (برحسب میکروپاسکال) یک صدا که بلندی آن ۶ سون است، چه مقدار است؟ سطح فشار صوتی صدایی که بلندی آن به لحاظ ذهنی دو برابر یک صدای ۶ سونی است، چه مقدار است؟
۲. با توجه به شکل ۴-۳، کدام یک به لحاظ ذهنی بلندتر است: یک موج سینوسی ۱۰۰ هرتزی با سطح فشار صوتی ۶۵ دسی بل یا یک موج سینوسی ۱۰۰۰ هرتزی با سطح فشار صوتی ۶۰ دسی بل؟ یک موج سینوسی ۱۵۰۰۰ هرتزی با سطح فشار صوتی ۸۰ دسی بل یا یک موج سینوسی ۲۰۰۰ هرتزی با سطح فشار صوتی ۷۰ دسی بل؟ یک موج سینوسی ۵۰۰ هرتزی با سطح احساس ۶۰ دسی بل یا یک موج سینوسی ۹۰۰۰ هرتزی با سطح احساس ۶۰ دسی بل؟
۳. با توجه به شکل ۴-۴، فرکانس شنیداری هر یک از موج‌های سینوسی زیر برحسب بارک چه مقدار است: ۷ کیلوهرتز، ۸ کیلوهرتز، ۱ کیلوهرتز، ۲ کیلوهرتز؟ افزایش کدام یک تفاوت بیشتری در فرکانس شنیداری را نشان می‌دهد؟ ۷ کیلوهرتز به ۸ کیلوهرتز یا ۱ کیلوهرتز به ۲ کیلوهرتز؟
۴. نمودار رابطه بین بلندی ذهنی با مسافت طی شده به وسیله صدا را رسم کنید. برای این کار می‌توانید مجموعه‌ای از صداها را زیر را به کار ببرید: صدای کارتی (مثلاً یک کارت ویزیت یا کارت اعتباری) که روی میز یا کاپوت ماشین ضربه می‌زند، صدای قلمی که روی میز ضربه می‌زند، صدای خط‌کشی

که روی میز ضربه می‌زند و صدای لیوان قهوه‌خوری که بر میز ضربه می‌زند. منظور این است که بلندی صداها در محدوده قابل کنترل باشد. می‌توانید از صداهای دیگری نیز استفاده کنید، اما دست‌کم چهار مورد انتخاب کنید و مواردی را انتخاب کنید که به لحاظ زمانی شبیه هستند (مانند ضربه‌زدن). این تمرین را با فرد دیگری انجام دهید. ابتدا، هر یک از شما به‌طور جداگانه، مقدار بلندی صدا را به‌صورت یک عدد در جاهای خالی زیر بنویسید (از کسر هم می‌توانید استفاده کنید).

(الف) قلم برابر بلندتر از کارت است.

(ب) خط‌کش برابر بلندتر از قلم است.

(ج) لیوان برابر بلندتر از خط‌کش است.

می‌توانید یک مقیاس بلندی ذهنی با گذاشتن عدد ۱ برای بلندی صدای کارت ایجاد کنید. سپس بلندی صدای قلم را ۱ برابر عددی که به (الف) داده‌اید، و بلندی صدای خط‌کش را برابر با بلندی قلم ضربدر عدد ارائه‌شده در (ب)، و بلندی لیوان را برابر با بلندی خط‌کش ضربدر عدد ارائه‌شده در (ج) قرار دهید. چقدر مقیاس شما با مقیاس نفر دیگر متفاوت است؟

حالا بلندی را به‌شکل دیگری اندازه بگیرید. یک راهرو یا یک حیاط یا یک دشت باز پیدا کنید. وقتی یکی از افراد صدایی تولید می‌کند، شخص دیگر باید آنقدر دور شود تا جایی که دیگر صدای فرد اول را نشنود. فردی که صدا تولید می‌کند باید تعداد قدم‌های شنونده را بشمارد و یادداشت کند. این کار باید برای هر صدا و برای هر دو نفر تکرار شود. مقیاس شما چقدر با مقیاس فرد دیگر متفاوت است؟

چقدر این دو مقیاس بلندی صدا به یکدیگر شباهت دارند؟ سعی کنید نمودار آنها را همانند نمودار سون در شکل ۴-۲ که برحسب فشار رسم شده است، رسم کنید. رابطه خطی است؟ یا غیرخطی؟ آیا یک مقیاس شباهت بیشتری به مقیاس سون و مقیاس دیگر شباهت بیشتری به مقیاس فشار صوتی دارد؟ چگونه مقیاس مسافت را به فشار صوتی تبدیل می‌کنید؟

فصل ۵

درک گفتار

وقتی به صحبت‌های کسی گوش می‌دهید، معمولاً بر فهم معنای گفتار او تمرکز می‌کنید. یکی از معروف‌ترین نقل‌قول‌ها در این زمینه (در زبان‌شناسی) این است: «ما حرف می‌زنیم برای آنکه شنیده شویم، برای آنکه فهمیده شویم» (یاکوبسن و همکاران^۱ ۱۹۵۲). تمایل ما به عنوان شنونده برای فهم گفتار گوینده، سبب می‌شود که بر معنای کلمات ادا شده تمرکز کنیم تا بر تلفظ دقیق آنها. اما بعضی اوقات شیوه تلفظ یک کلمه توجه شما را جلب می‌کند: فردی یک کلمه آشنا را به گونه‌ای ناآشنا بیان می‌کند و شما بلافاصله می‌پرسید «شما این کلمه را این طور تلفظ می‌کنید؟» هنگامی که به آواشناسی گفتار گوش می‌دهیم - به اینکه کلمات چگونه بیان می‌شوند و نه اینکه معنای آنها چیست - به عنوان شنونده با درک گفتار سروکار داریم.

در درک گفتار، شنوندگان توجه خود را به صداها و گفتاری معطوف می‌کنند و به جزئیات آوایی در تلفظ اهمیت می‌دهند که اغلب در ارتباط گفتاری معمولی اصلاً مورد توجه قرار نمی‌گیرند. به عنوان مثال، شنوندگان اغلب در محاوره معمولی یک

1. Jakobson et al.

خطای گفتاری یا یک اشتباه تلفظی عمدی را نمی‌شنوند و یا به نظر نمی‌رسد که بشنوند، اما هنگامی که برای شنیدن اشتباهات تلفظی آموزش می‌بینند، متوجه همین خطاها می‌شوند (رجوع شود به کول^۱ ۱۹۷۳).

آزمایش تشخیص اشتباه تلفظی

همان‌طور که مشغول امور روزمره خود هستید، سعی کنید گاه‌گاه یک کلمه را اشتباه تلفظ کنید و ببینید آیا افرادی که با آنها مشغول صحبت هستید، متوجه آن می‌شوند یا نه. به‌عنوان مثال، اگر درباره کلاس بیولوژی صحبت می‌کنید، این کلمه را به‌صورت «بیولوچی» تلفظ کنید. پس از آنکه این کلمه را یک یا دو بار بدین‌صورت تلفظ کردید، درباره آزمایش خود به دوستانتان توضیح دهید و از آنها پرسید که آیا متوجه هیچ اشتباه تلفظی شدند یا نه. آیا افراد بیشتر متوجه اشتباه تلفظی در جایگاه آغاز کلمه می‌شوند یا در میان آن؟ در واژه‌ها بیشتر متوجه اشتباه تلفظی می‌شوند یا در همخوان‌ها در اسم‌ها و افعال یا در کلمات دستوری؟ افراد اگر متوجه تلفظ اشتباهی یک کلمه نمی‌شوند، چگونه کلمات را در فرهنگ لغات ذهنی‌شان پیدا می‌کنند؟ واضح است که جستجوی کلمات در واژگان ذهنی تا حدی متفاوت از جستجوی کلمات در یک فرهنگ لغت چاپی است (سعی کنید «بیولوچی» را در گوگل جستجو کنید). آیا فکر می‌کنید وقتی از روی عمد بر تلفظ اشتباهی کلمات اصرار می‌ورزید، دوستانان شما را غیرعادی تصور می‌کنند؟

بدین ترتیب در این فصل، درک گفتار را به‌عنوان جنبه آوایی شنیدن مورد بحث قرار می‌دهیم و در آن نه بر کلمات، بلکه بر صداها و گفتاری تمرکز می‌کنیم. یکی از مسائل جالب در آواشناسی و روان‌شناسی زبان یافتن روشی برای اندازه‌گیری میزان اطلاعات آوایی است که شنوندگان در طی محاوره عادی دریافت می‌کنند، اما در این کتاب تمرکز ما منحصر به جنبه آوایی شنیدن است.

۵-۱. توانایی شنیداری درک گفتار را شکل می‌دهد

همان‌طور که در فصل ۴ دیدیم، درک گفتار از طریق ویژگی‌های عام سیستم شنوایی شکل می‌گیرد که تعیین می‌کند چه چیز می‌تواند شنیده شود و چه چیز نمی‌تواند، چه سرنخ‌هایی^۱ در بافت‌های آوایی خاص قابل‌بازیابی‌اند و چگونه صداهای مجاور بر یکدیگر تأثیر می‌گذارند. برای مثال، دیدیم که احتمالاً غیرخطی بودن مقیاس فرکانسی حلزون گوش علت این واقعیت است که در هیچ زبانی تمایز همخوان‌های سایشی بر اساس مؤلفه‌های فرکانسی بالای ۶۰۰۰ هرتز صورت نمی‌گیرد.

دو مثال دیگر نیز نشان می‌دهد که چگونه سیستم شنوایی، درک گفتار را محدود می‌کند. مثال اول مربوط به تفاوت بین انسدادی‌های دمیده^۲ و نادمیده است. این تقابل به‌وسیله یک سرنخ زمانی که «زمان آغاز واک» (وی‌آتی)^۳ نامیده می‌شود، در سیگنال گفتار انتقال پیدا می‌کند. وی‌آتی مدت‌زمان تأخیر شروع واک (برحسب میلی‌ثانیه) پس از انفجار رهش یک انسدادی است. مدت‌زمان تأخیر شروع واک در انسدادی‌های دمیده طولانی‌تر از انسدادی‌های نادمیده است، زیرا در انسدادی‌های دمیده، تارهای صوتی برای مدت کوتاهی پس از آنکه گرفتگی دهانی در همخوان انسدادی رهش یافت، باز می‌مانند و بدین‌شکل، فوت کوتاه هوا در انسدادی‌های دمیده بی‌واک تولید می‌شود. مشاهدات نشان می‌دهد که در بسیاری از زبان‌ها، بین انسدادی‌های دمیده و نادمیده مرزی در حدود ۳۰ میلی‌ثانیه وی‌آتی وجود دارد. یک تأخیر ۳۰ میلی‌ثانیه‌ای بین رهش انسدادی و آغاز واک چه ویژگی خاصی دارد؟

در اینجا است که نقش سیستم شنوایی روشن می‌شود. توانایی ما به‌عنوان شنونده در تشخیص آغازهای^۴ ناهم‌زمان نواخت‌ها در فرکانس‌های مختلف، احتمالاً علت اصلی این امر است که رایج‌ترین مرز زمان آغاز واک در زبان‌ها در حدود ± 30 میلی‌ثانیه است. دو نواخت خالص، یکی در ۵۰۰ هرتز و دیگری در ۱۰۰۰ هرتز را در نظر بگیرید. در یک آزمایش درک (رجوع کنید به مطالعات پژوهشی پیسونی^۵

1. cue

2. aspirated

3. voice onset time (VOT)

4. onset

5. Pisoni

۱۹۷۷، پاستور^۱ و فرینگتون^۲ (۱۹۹۶) ما این دو نواخت را با اندکی ناهم‌زمانی آغاز ترکیب کردیم، به نحوی که نواخت ۵۰۰ هرتزی ۲۰ میلی‌ثانیه زودتر از نواخت ۱۰۰۰ هرتزی آغاز شد. وقتی از شنندگان خواسته شد، قضاوت کنند که آیا دو نواخت هم‌زمان آغاز می‌شود یا یکی اندکی زودتر از دیگری، دریافتیم که شنندگان تصور می‌کنند که نواخت‌های جداشده با ۲۰ میلی‌ثانیه ناهم‌زمانی آغاز، هم‌زمان آغاز می‌شوند. شنندگان متوجه ناهم‌زمانی آغاز نمی‌شوند تا اینکه اختلاف به حدود ۳۰ میلی‌ثانیه برسد. این شباهت بین درک شنیداری غیرگفتاری و یک جهانی آوایی [وی‌آئی] منجر به طرح این دیدگاه شد که توانایی سیستم شنوایی در تشخیص ناهم‌زمانی آغاز، احتمالاً عامل اصلی این ویژگی آوایی جهانی است.

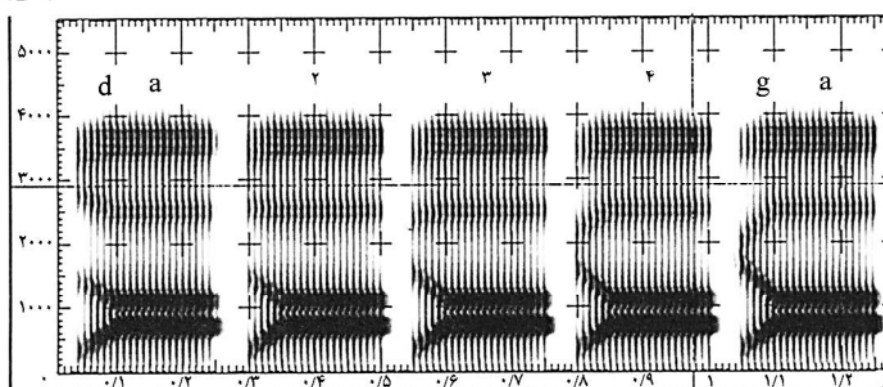
مثال شماره دو: دیگر ویژگی عام سیستم شنوایی، احتمالاً در پدیده درکی به نام «جبران هم‌تولیدی»^۳ دخیل است. این اثر در درک محل تولید در هجاهای CV روی می‌دهد. ابزار اصلی برای مطالعه این اثر، پیوستار هجایی است که با گام‌های آکوستیکی یکسان از [da] تا [ga] تغییر می‌کند (شکل ۵-۱). این شکل اندکی توضیح نیاز دارد. در پایان فصل ۳، طیف‌نگاشت معرفی و به این نکته اشاره شد که نوارهای سیاه طیف‌نگاشت، قله‌های طیفی را نشان می‌دهند که به وسیله بازخوان‌های دستگاه گفتار تولید می‌شوند (فرکانس‌های سازه‌ای). از این رو، در شکل ۵-۱ (الف) توالی از ۵ هجا را می‌بینیم که در آن هجای شماره ۱ با عنوان [da] و هجای شماره ۵ با عنوان [ga] مشخص شده است. در همه هجاها، واکه یکسان است؛ فرکانس سازه اول آن (F_1) در حدود ۹۰۰ هرتز، فرکانس سازه دوم (F_2) در حدود ۱۱۰۰ هرتز، F_3 در حدود ۲۵۰۰ هرتز و F_4 در حدود ۳۷۰۰ هرتز است. تفاوت بین هجاهای [da] و

1. Patore 2. Farrington

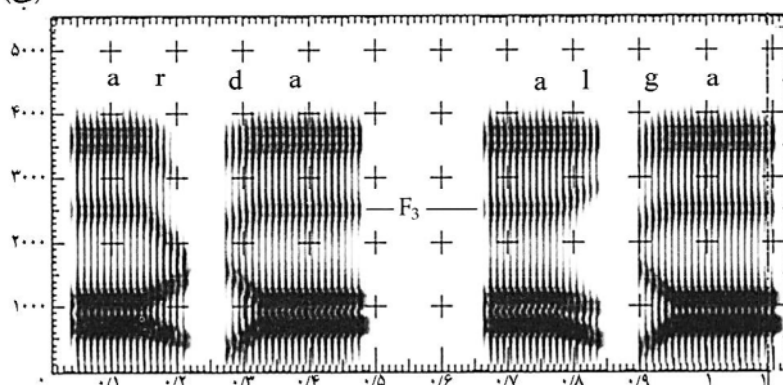
۳. compensation for coarticulation. هنگام تولید گفتار، فرایند هم‌تولیدی سبب می‌شود مشخصات یک آوا تحت تأثیر آواهای مجاور قرار گیرد، اما شنونده در درک گفتار با حذف تأثیرات بافتی و بازایی مشخصات آوایی زیرساختی هر واج، هم‌تولیدی را جبران می‌کند (Elman and McClelland 1988). (مترجم)

[ga] ناشی از حرکت‌های کوتاه سازه‌ای (به‌نام گذرهای سازه‌ای^۱) در آغاز هر هجاست. در هجای [da] فرکانس F_2 در ۱۵۰۰ هرتز و فرکانس F_3 در ۲۹۰۰ هرتز شروع می‌شود، در حالی که در هجای [ga] فرکانس F_2 در ۱۹۰۰ هرتز و فرکانس F_3 در ۲۰۰۰ هرتز آغاز می‌شود. در شکل ۵-۱ (ب) می‌بینید که تفاوت اصلی بین هجاهای [ar] و [al] در الگوی F_3 در پایان هجاست.

(الف)



(ب)



شکل ۵-۱. (الف) پیوستار هجاهای سنتز شده همخوان-واکه از "da" تا "ga" در پنج گام آکوستیکی یکسان. (ب) اگر هجای ماقبل نمونه آوایی شماره ۳ در پیوستار "da/ga"، "ar" باشد، صدایی شبیه به "da" دارد و اگر «al» باشد، صدایی شبیه به "ga" دارد.

1. formant transition

ویرجینیا مان^۱ (۱۹۸۰) به این نتیجه رسید که درک پیوستار [da]-[ga] به بافت پیشین بستگی دارد. پاسخ شنوندگان در آزمایش وی نشان داد که هجاهای مبهم در میانه این پیوستار اگر هجای VC ماقبل آنها [al] باشد، صدایی شبیه به "ga" دارند و اگر هجای ماقبل آنها [ar] باشد، صدایی شبیه به "da" دارند.

پدیده درکی «جبران هم‌تولیدی» همان‌طور که عنوان آن نشان می‌دهد، می‌تواند مربوط به هم‌تولیدی بین همخوان پایانی در نمونه بافتی VC ([ar] یا [al]) و همخوان آغازین در نمونه آزمایشی CV ([da]-[ga]) باشد. اما اثر تقابلی^۲ یک فرکانس شنیداری نیز احتمالاً نقشی ایفا می‌کند. این تبیین در شکل ۵-۱ (ب) نشان داده شده است. فرکانس نسبی F_3 هجای [da] را از [ga] متمایز می‌سازد، فرکانس F_3 در هجای [da] بالاتر از [ga] است. در عین حال، جالب است که فرکانس درکی F_3 ممکن است تحت تأثیر فرکانس F_3 هجای ماقبل [da/ga] نیز باشد. هرگاه فرکانس F_3 درست قبل از [da/ga] پایین باشد (همچون در [ar])، فرکانس F_3 هجای [da/ga] به‌طور تقابلی بالاتر به نظر می‌رسد و هرگاه F_3 درست قبل از [da/ga] بالا باشد، فرکانس F_3 نمونه [da/ga] به‌طور تقابلی پایین‌تر به نظر می‌رسد. لوتو و کلوندر^۳ (۱۹۹۸) این یافته را با جایگزینی هجای پیشرو با یک موج سینوسی ساده که در یک حالت با فرکانس پایانی F_3 در هجای [ar] و در حالت دیگر با فرکانس پایانی F_3 در هجای [al] مطابقت داشت، آزمایش کردند. آنها دریافتند که این نواخت‌های مجزای غیرگفتاری، درک پیوستار [da]-[ga] را در همان جهتی تغییر می‌دهند که هجاهای [ar] و [al] تغییر دادند. از این رو، واضح است که حداقل بخشی از پدیده جبران هم‌تولیدی ناشی از اثر یک تقابل شنیداری ساده است که ربطی به جنبه آواشناختی درک ندارد.

1. Virginia Mann

2. contrast effect

3. Lotto and Kluender

دو تبیین برای یک اثر

جبران هم‌تولیدی، موضوع بحث‌برانگیزی است. برای پژوهشگرانی که تمایل دارند درک گفتار را در چارچوب درک آوایی در نظر بگیرند، یعنی این افراد «شنوا» هستند که حرف می‌زنند، جبران هم‌تولیدی بر اساس هم‌تولیدی تبیین می‌شود. عقب‌کشیدن زبان در [r] سبب می‌شود که شنوندگان عقب‌کشیده‌شدن زبان در آوای بعدی را نیز پیش‌بینی کنند و در نتیجه، یک انسدادی پسین (که بیشتر شبیه به «g» است) ممکن است در بافت [r]، به دلیل این پیش‌بینی وابسته به بافت، همچنان بیشتر شبیه به «d» به نظر برسد. پژوهشگرانی که معتقدند برای تبیین اثرات درکی باید پیش از هر چیز، سیستم درون‌داد حسی را مورد توجه قرار داد (پیش از مفروض دانستن تبیین‌های انتزاعی‌تر تقطیع شناختی)، کاملاً تحت تأثیر دیدگاه تقابل شنیداری قرار دارند.

به نظر من شواهد نشان می‌دهد که هر دوی این تبیین‌ها درست هستند. تقابل شنیداری به نظر می‌رسد در نمونه‌هایی که بافت آنها به جای [ar] یا [al]، نواخت خالص است، نیز روی می‌دهد، اما میزان اثر آن در مقایسه با بافت یک هجای پیشرو آوایی کمتر است. کم‌تر بودن میزان اثر تقابل شنیداری نشان می‌دهد که این اثر یگانه عامل نیست. من نیز تحقیقی با همین محرک‌ها^۱ با تهیه پیوستاری بین [al] و [ar] به عنوان بافت برای پیوستار [da]-[ga] انجام دادم. وقتی که هم هجای پیشرو و هم هجای هدف مبهم هستند، ماهیت هجای هدف (به عنوان “da” یا “ga”) به ماهیت درکی هجای پیشرو بستگی دارد. به عبارت دیگر، برای یک نمونه آکوستیکی واحد، اگر شنونده تصور کند که بافت پیشرو “ar” است، احتمال بیشتری دارد که هجای مبهم را “da” شناسایی کند. این مطمئناً اثر یک تقابل شنیداری نیست.

از این رو، به نظر می‌رسد، هم درک شنیداری و هم درک آوایی، شنونده را به یک سو هدایت می‌کنند.

۵-۲. دانش آوایی درک گفتار را شکل می‌دهد

البته این واقعیت که سیستم شنوایی، درک ما از گفتار را شکل می‌دهد، بدین معنا نیست که همه فرایندهای درک گفتار به وسیله قابلیت‌های شنیداری ما صورت می‌گیرند. ما به عنوان گویشوران زبان و نه صرفاً شنوندگان آن، به وسیله دانشمان از تولید گفتار نیز هدایت می‌شویم. دو نوع اصلی اثرات درکی که منتج از دانش آوایی هستند، عبارت‌اند از: درک مقوله‌ای^۱ و انسجام آوایی^۲.

۵-۲-۱. درک مقوله‌ای

به شکل ۵-۱ (الف) دوباره نگاهی بیاندازید. در این شکل توالی از هجاها دیده می‌شود که به تدریج (و در گام‌های آکوستیکی مساوی) از هجایی شبیه به “da” در یک انتهای پیوستار به هجایی شبیه به “ga” در انتهای دیگر پیوستار تغییر می‌کنند (جدول ۵-۱ را ببینید). این نوع توالی که به تدریج تغییر می‌کند، یک پیوستار محرکی نامیده می‌شود. هنگامی که این هجاهای سنتز شده را برای شنوندگان پخش می‌کنیم و از آنها می‌خواهیم با دستورالعملی مثل «لطفاً آنچه را که می‌شنوید، بنویسید»، آنها را شناسایی کنند، افراد معمولاً سه هجای اول را “da” و دو هجای آخر را “ga” می‌نامند. پاسخ آنها بسیار مقوله‌ای به نظر می‌رسد: یک هجا یا “da” است یا “ga”. البته این خیلی ساده است، زیرا تنها دو برجسب برای صداهای پیوستار داریم و بنابراین، بر اساس تعریف، افراد باید بگویند یا “da” یا “ga”. با این وجود، جالب است که توانایی شنیدن تفاوت بین محرک‌های پیوستار بر اساس برجسب‌هایی که ما برای شناسایی آنها به کار می‌بریم، قابل پیش‌بینی است و به همین دلیل است که می‌گوییم درک گفتار گرایش مقوله‌ای دارد.

جدول ۵-۱. پارامترهای مهم آکوستیکی و نتایج آزمون شناسایی هجاهای شکل ۵-۱ (الف).

شناسایی شده به صورت	آغازۀ F ₃	آغازۀ F ₂	شمارۀ نمونه
“da”	۲۷۵۰	۱۴۸۰	۱
“da”	۲۵۶۲	۱۵۲۲	۲
“da”	۲۳۷۵	۱۵۶۵	۳
“ga”	۲۱۸۷	۱۶۰۷	۴
“ga”	۲۰۰۰	۱۶۵۰	۵

1. categorical perception

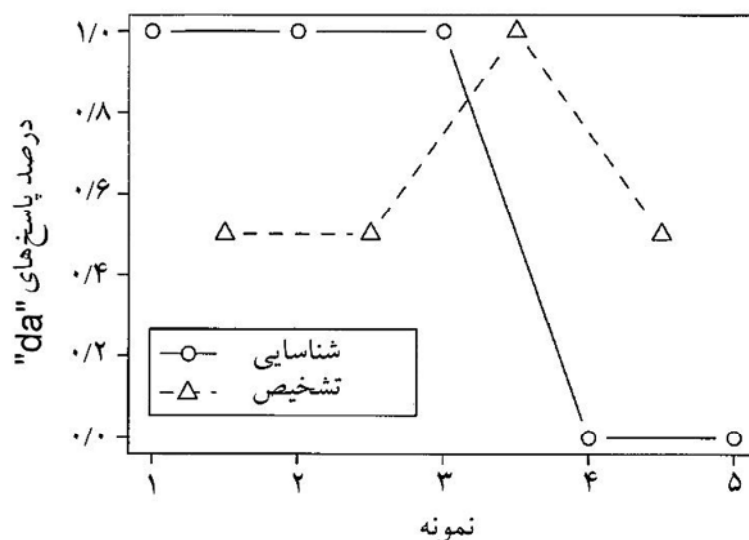
2. phonetic coherence

برای روشن شدن این مطلب، فرض کنید که من دو هجای اول در پیوستار شکل ۵-۱، نمونه های شماره ۱ و ۲ را برای شما پخش کنم. شنوندگان به هر دوی این نمونه ها برچسب “da” می دهند، اما این نمونه ها تفاوت کمی با یکدیگر دارند. آغاز سازۀ سوم در نمونه شماره ۱ در ۲۷۵۰ هرتز است، در حالی که F_3 در نمونه شماره ۲ در ۲۵۶۲ هرتز آغاز می شود. افراد متوجه این تفاوت نمی شوند، دو هجا گویی صدایی واقعاً شبیه به یکدیگر دارند. همین وضعیت با مقایسه نمونه ۲ با ۳ و نمونه ۴ با ۵ نیز ملاحظه می شود. اما وقتی که نمونه ۳ را می شنوید (هجایی که معمولاً به آن برچسب “da” می دهید) در مقایسه با نمونه ۴ (هجایی که معمولاً به آن برچسب “ga” می دهید)، تفاوت بین آنها بلادرنگ برایتان مشخص می شود. نکته این است که در وظیفۀ تشخیص^۱ وقتی از شما خواسته می شود که تفاوت های کوچک را تشخیص دهید، نباید از برچسب های “da” یا “ga” استفاده کنید. شما باید قادر به شنیدن تفاوت ها دقیقاً در همان سطح دقت باشید، صرف نظر از اینکه نمونه ها چه برچسبی دارند، زیرا اختلاف (۱۸۸ هرتز در آغازۀ F_3) برای نمونه ۱ در مقابل ۲ به همان میزانی است که برای نمونه ۳ در مقابل ۴ است. واقعیت شگفت انگیز این است که حتی وقتی نباید از برچسب های “da” و “ga” در پاسخ های شنیداری خود استفاده کنید، درک شما با برچسب ها مطابقت دارد - می توانید وقتی که نمونه ها برچسب های مختلفی دارند، متوجه یک اختلاف ۱۸۸ هرتزی بشوید، اما وقتی که نمونه ها برچسب های یکسانی دارند، خیلی متوجه نمی شوید.

یک روش کلاسیک برای نمایش این نتایج نظری در شکل ۵-۲، نشان داده شده است (برای مشاهده نمودارهایی از این دست رجوع کنید به لیبرمن و همکاران ۱۹۵۷). این نمودار دو «تابع» دارد: دو خط، یکی برای نمایش نسبت دفعاتی که شنوندگان، یک نمونه را به عنوان “da” شناسایی کرده اند و دیگری برای نمایش نسبت دفعاتی که شنوندگان می توانند دقیقاً بگویند که آیا دو نمونه (به عنوان مثال شماره ۱ و شماره ۲) با

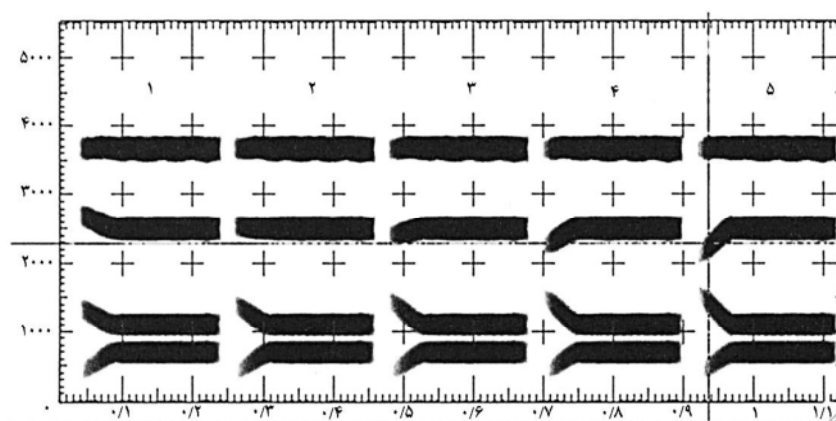
1. discrimination task

یکدیگر تفاوت دارند یا خیر. تابع اول، تابع شناسایی^۱ نامیده می‌شود و من آن را به گونه‌ای رسم کردم که انگار ما همیشه (با احتمال مساوی ۱) نمونه‌های ۱ و ۲ و ۳ را به صورت "da" شناسایی می‌کنیم. تابع دوم، تابع تشخیص نامیده می‌شود و من موردی را رسم کردم که نشان می‌دهد حدس شنونده درباره نمونه‌هایی که برچسب یکسانی دارند، کاهش یافته است (مقدار «حدس» که برابر با احتمال تشخیص درست تفاوت است، مساوی ۰/۵ می‌باشد) و نیز نشان می‌دهد که شنونده همواره قادر بوده که تفاوت بین نمونه ۳ (با برچسب "da") و نمونه ۴ (با برچسب "ga") را بشنود. الگوی پاسخ در شکل ۲-۵، همان چیزی است که منظور ما از «درک مقوله‌ای» است، تشخیص درون مقوله‌ای شانس است، اما تشخیص بین مقوله‌ای کامل. گفتار تمایل دارد که به صورت مقوله‌ای درک شود، هرچند که مانند جبران هم‌تولیدی، یک مؤلفه درک شنیداری نیز در این نوع آزمایش وجود دارد، بنابراین، درک گفتار هرگز به طور کامل مقوله‌ای نیست.



شکل ۲-۵. الگوی کلاسیک بازنمایی درک مقوله‌ای. پاسخ‌های شناسایی (با خطی پیوسته و دایره) گذری تند از یک مقوله به مقوله دیگر را نشان می‌دهند و پاسخ‌های تشخیص (یعنی احتمال پاسخ «متفاوت» با خط شکسته و مثلث) نشانگر چیزی بیشتر از تشخیص درون مقوله‌ای شانی نیستند.

گرایش ما برای درک مقوله‌ای گفتار به شیوه‌های بسیار مختلفی مورد بررسی قرار گرفته است. یکی از جالب‌ترین این روندهای پژوهشی (حداقل برای من) نشان می‌دهد که درک مقوله‌ای گفتار یک پدیده آموختنی است (رجوع کنید به جانسون و رالستون^۱ ۱۹۹۴). این پژوهش‌ها نشان می‌دهد که درک امواج سینوسی آنالوگ در پیوستار “da” به “ga” در مقایسه با درک گفتار طبیعی خیلی کمتر مقوله‌ای است. رابرت رمز^۲ و همکاران (رمز و همکاران ۱۹۸۱) پیشنهاد استفاده از امواج سینوسی آنالوگ گفتاری برای مطالعه درک گفتار هستند. در امواج سینوسی آنالوگ، به جای سازه‌ها، امواج سینوسی متغیربازمان^۳ جایگزین می‌شود (شکل ۵-۳). این سیگنال‌ها هرچند که به لحاظ آکوستیکی شبیه به گفتار هستند، هرگز صدایی شبیه به گفتار طبیعی ندارند. این واقعیت که ما پاسخ مقوله‌ای‌تر به سیگنال‌های گفتاری داریم تا به امواج سینوسی آنالوگ، نشان می‌دهد که ویژگی خاصی در شنیدن فرکانس‌های سازه‌ای به صورت گفتار در مقایسه با شنیدن آنها به صورت نوبه‌های غیرگفتاری بازی‌های ویدیویی وجود دارد. یک تبیین این است که ما به عنوان انسان دارای توانایی ذاتی برای بازیابی اطلاعات آوایی سیگنال گفتار هستیم، به نحوی که ما بستر مقوله‌ای مورد نظر گوینده را می‌شنویم.



شکل ۵-۳. پیوستار هجاهایی که با امواج سینوسی آنالوگ و طبق پیوستار “da-ga” در شکل ۵-۱ مدل‌سازی شده‌اند.

1. Johnson and Ralston

2. Robert Remez

3. time-varying

تبیین مشابه دیگری در مورد اینکه چرا درک گفتار گرایش مقوله‌ای دارد، این است که سیستم‌های درکی ما بر اساس تجربه زبانی ما تنظیم می‌شوند. ما به‌عنوان گویشوران وقتی حرف می‌زنیم، تاحدی نیت مقوله‌ای داریم، مثلاً با گفتن "dot" به جای "got". از این رو، به‌عنوان شنونده، گفتار را طبق مقوله‌هایی ارزیابی می‌کنیم که آموخته‌ایم به‌عنوان گوینده به کار ببریم. شواهد متعددی این دیدگاه «مقوله‌بندی اکتسابی» در درک مقوله‌ای را تأیید می‌کند.

به‌طور مثال، وقتی تلاش می‌کنید آواهای الفبای بین‌المللی آوایی را یاد بگیرید، ملاحظه کرده‌اید که آواهای گفتاری خارجی اغلب شبیه به آواهای زبان مادری شنیده می‌شوند. به‌عنوان نمونه، برای بیشتر افراد وقتی شروع به یادگرفتن صداهای مکیده [b]، [d]، و [g] می‌کنند، شنیدن تمایز آنها با انسدادی‌های واکدار ساده دشوار است. این مشاهده ساده، بارها و به‌شیوه‌های متعدد تأیید شده است و نشان می‌دهد که ما در درک گفتار، صداهایی را می‌شنویم که به‌عنوان گویشور با آنها آشنا هستیم. مرزهای درک مقوله‌ای ما به‌وسیله زبانی که بدان صحبت می‌کنیم، تعیین می‌شود. (به نظریه‌های بست^۱ ۱۹۹۵ و فلج^۲ ۱۹۹۵، رجوع کنید).

آهن‌رباهای مقوله‌ای^۳

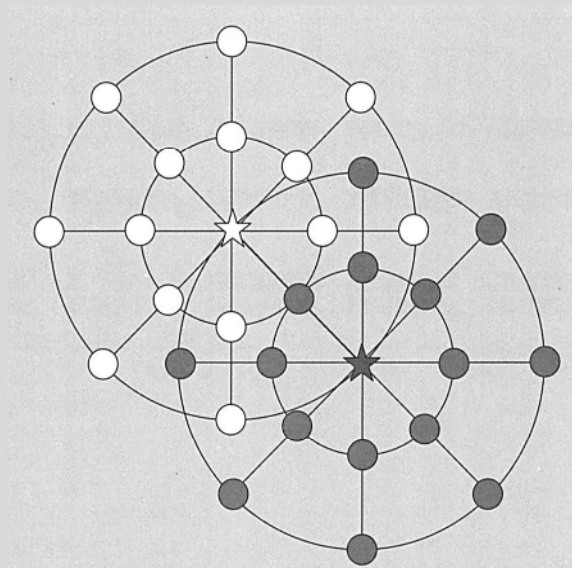
مثال واقعاً جالبی که نشان می‌دهد درک مقوله‌ای زبان‌ویژه است، «اثر آهن‌ربای درکی^۴» است (کوهل^۵ و همکاران ۱۹۹۲). در این آزمایش، یک واکه که نمونه متعارف صدای [i] است، سنتز می‌شود و سپس در اطراف آن واکه‌هایی که به‌طور منظم با واکه مرکزی تفاوت دارند، قرار می‌گیرند. در شکل ۴-۵، این آزمایش با یک ستاره سفید و دایره‌های سفید اطراف آن نشان داده شده است. دسته دومی از واکه‌ها

1. Best 2. Flege 3. categorical magnet 4. perceptual magnet effect

اثر آهن‌ربای درکی سبب می‌شود که فاصله درکی بین یک نمونه اصلی آوایی با گونه‌های اطراف آن کاهش یافته و یا از بین برود، زیرا آوای اصلی همانند یک آهن‌ربا گونه‌های مجاور خود را جذب می‌کند (کوهل و ایورسون ۱۹۹۵، ۱۲۳). (مترجم)

5. Kuhl et al.

دوباره به صورت یک شبکه شعاعی در اطراف واژه مرکزی ستنز می شود. در مرکز دسته دوم، نمونه متعارف واژه [i] قرار نمی گیرد، بلکه واژه ای قرار می گیرد که به مرز بین [i] و [e] نزدیک تر است.

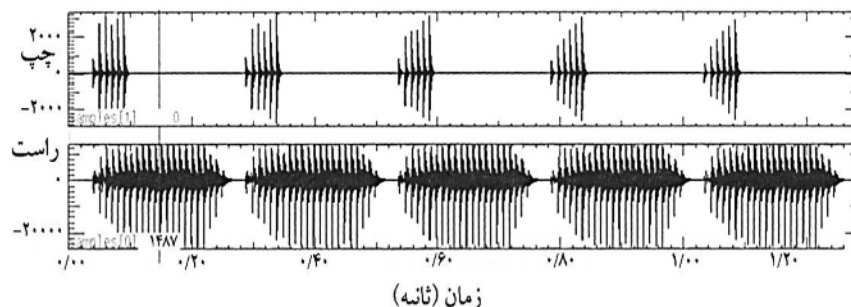


شکل ۵-۴. مدل کلی محرک هایی که برای مقایسه حساسیت درکی حول نمونه متعارف یک واژه (ستاره سفید) با حساسیت درکی حول نمونه نامتعارف همان واژه به کار رفتند.

وقتی از بزرگسالان می پرسید که آیا می توانند تفاوت بین واژه مرکزی (یکی از ستاره ها) و اولین حلقه واژه ها را بشنوند، به نظر می رسد که برای تشخیص ستاره سفید (نمونه متعارف [i]) از واژه های اطرافش در مقایسه با تشخیص ستاره سیاه (نمونه غیرمتعارف [i]) از نمونه های مجاورش با دشواری بیشتری مواجه هستند. این اثر، جالب است، زیرا نشان می دهد که درک مقوله ای در درون مقوله ها مدرج است (لازم به ذکر است که همه واژه های آزمایش صدایی شبیه به واج گونه های [i] دارند، حتی واژه هایی که در دسته سیاه قرار دارند و نزدیک به مرز [i]/[e] هستند). با این همه، حتی جالب تر از آن، این واقعیت است که مکان یک آهن ربای درکی بسته به زبان مادری شنونده متفاوت است، حتی وقتی این شنندگان، کودکانی بیش نیستند.

۵-۲-۲. انسجام آوایی^۱

تجربه حسی شنیداری به وسیله شماری از اصول نظم‌دهنده گشتالت^۲ که «تجزیه صحنه شنیداری»^۳ نامیده می‌شوند، «تصویری» منسجم از جهان پدید می‌آورد (برگمن^۴ ۱۹۹۰). وقتی گفتار را درک می‌کنیم، انسجام آوایی را با مؤلفه‌های آکوستیکی تجربه می‌کنیم که طبق اصول تجزیه صحنه باید نامنسجم باشند.



شکل ۵-۵. شکل موج محرک‌هایی که برای آزمایش درک دوطرفه به کار رفتند. قسمت بالای شکل، سیگنالی را نشان می‌دهد که به گوش چپ فرستاده می‌شود و قسمت پایین شکل، سیگنالی را که به گوش راست می‌رود، نشان می‌دهد.

۱. منظور از انسجام آوایی این است که مؤلفه‌های یک صدا، به‌عنوان مثال صدای یک ویلن در یک ارکستر، مجموعه تقریباً منسجمی را می‌سازند که صدای آن را از سایر صداها مثلاً صدای گیتار و تنبک متمایز می‌کند. (مترجم)

2. gestalt

بر اساس نظریه روان‌شناسی گشتالت، عناصر حسی (مثلاً شنیداری) بر پایه برخی اصول در مغز سازماندهی شده و الگوهای واحدی را شکل می‌دهند. مثلاً بر اساس یکی از اصول آن: عناصر هم‌جوار، هم‌شکل، هم‌اندازه به یکدیگر وصل شده، یک کل را به وجود می‌آورند (برگمن ۱۹۹۰). (مترجم)

3. auditory scene analysis

در موقعیت‌های واقعی زندگی، گوش در معرض صداهایی قرار می‌گیرد که از منبع‌های مختلف منتشر می‌شوند. این صداها در حوزه زمان و فرکانس با یکدیگر در هم آمیخته و به‌صورت یک موج فشار صوتی واحد وارد گوش می‌شوند. گوش چگونه این صداها را از یکدیگر تفکیک می‌کند؟ به‌عنوان مثال، در یک خیابان شلوغ، چگونه صدای یک فرد خاص، متفاوت از صدای بوق ماشین‌ها، جمعیت، زوزه باد و غیره درک می‌شود. سیستم شنوایی طی فرایند «تجزیه صحنه شنیداری» ابتدا با تجزیه کردن سیگنال آکوستیکی ورودی و سپس ترکیب کردن اطلاعات شنیداری که از یک منبع منتشر شده‌اند (بر اساس اصول گشتالت)، صداها را از یکدیگر تفکیک می‌کند (برگمن ۱۹۹۰). (مترجم)

4. Bregman

درک دوطرفه^۱ مثال خوبی برای انسجام آوایی است. برای آزمایش این پدیده که توسط تیموتی راند^۲ در سال ۱۹۷۴ کشف شد، یک نوفه^۳ چهلچه‌ای^۴ کوتاه - یک نواخت غلتان^۵ ۸۰ میلی‌ثانیه‌ای کوتاه که متناظر با فرکانس معمول F₃ در یک هجای [da] یا [ga] است - به گوش چپ فرستاده می‌شود و به گوش راست، یک محرک «پایه»^۵ که از هجاهای [da] یا [ga] ترکیب یافته و فاقد مؤلفه^۶ چهلچه‌ای F₃ است، فرستاده می‌شود. جالب است که نمونه پایه می‌تواند برای [da] و [ga] یکسان باشد. از این رو، تنها تفاوت بین محرک‌ها در نوفه^۷ «چهلچه‌ای» است. شکل ۵-۵ شکل موج آکوستیکی ۵ هجایی را که به گوش چپ و راست فرستاده می‌شوند، نشان می‌دهد. اولین هجا صدایی شبیه به [da] دارد و آخرین هجا صدایی شبیه به [ga] (این توالی درست شبیه به پیوستار شکل ۵-۱ (الف) است). سیگنال پایه به گوش راست و نوفه‌های چهلچه‌ای به گوش چپ فرستاده می‌شود.

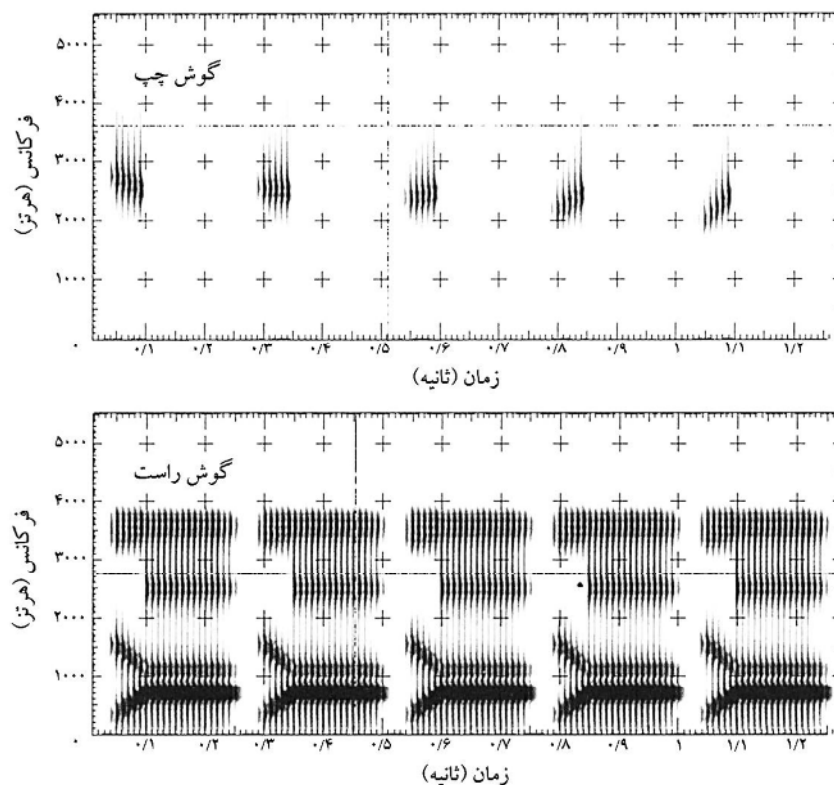
طیف‌نگاشت‌های شکل ۵-۶، ویژگی خاص محرک‌های درک دوطرفه را نشان می‌دهد. همان‌طور که در طیف‌نگاشت‌های پایین شکل ۵-۶ می‌بینید، محرک پایه در هر پنج هجا یکسان است و در هر هجا، یک فاصله خالی در جایی که سازه سوم باید باشد، دیده می‌شود. چهلچه‌هایی که در بالای شکل ۵-۶ دیده می‌شوند، این فواصل خالی را کاملاً پر می‌کنند. اولین هجا چهلچه‌ای رو به کاهش و آخرین هجا، چهلچه‌ای رو به افزایش دارد. وقتی چهلچه‌ها را به محرک‌های پایه می‌افزایید، تقریباً به‌طور کامل پیوستار [da]-[ga] که در شکل ۵-۱ (الف) نشان داده شد، به دست می‌آید.

1. duplex perception 2. Timothy Rand 3. chirp noise

چهلچه، سیگنالی است که فرکانس آن با زمان، افزایش یا کاهش می‌یابد و عنوان آن اشاره دارد به صدای چهلچه‌ای که پرندگان تولید می‌کنند. (مترجم)

۴. tone glide. نواخت غلتان، یک نواخت فرکانس متغیر است. (مترجم)

5. base stimulus



شکل ۵-۶. طیف نگاشت محرک‌های در که دوطرفه. محرک‌ها در واقع شبیه به محرک‌های شکل ۵-۱ هستند با این تفاوت که اطلاعات اصلی که [da] را از [ga] متمایز می‌کند، جدا شده و به گوش چپ فرستاده می‌شود.

در درک شنیداری معمولی، صداهاایی که در گوش چپ بلندتر احساس می‌شوند، به نظر می‌آید که از سمت چپ بدن می‌آیند. حال آنکه صداهاایی که در گوش راست بلندتر احساس می‌شوند، به نظر می‌آید از سمت راست بدن است. این در پدیده درک دوطرفه نیز صادق است. به نظر می‌رسد، چهچه در سمت چپ و نمونه پایه در سمت راست باشد. گرچه مسئله دیگری نیز روی می‌دهد که در درک شنیداری معمولی مشاهده نمی‌کنیم. چهچه (هرچند که از جهتی متفاوت نسبت به پایه به گوش می‌رسد) بر درک آوایی پایه تأثیر می‌گذارد. اگر چهچه شبیه به F_3 هجای [da] باشد، شنندگان هجای پایه را به صورت “da” می‌شنوند و اگر چهچه

شبیه به F_3 هجای [ga] باشد، شنوندگان هجای پایه را به صورت "ga" می‌شنوند. این پدیده «دوطرفه» نامیده می‌شود، زیرا به نظر می‌رسد چهچه در یک لحظه در دو محل است: به صورت یک نوفه چهچه‌ای غیرگفتاری مجزا و نیز به صورت یک مؤلفه آوایی تعیین‌کننده محل تولید هجای پایه. این یک اثر ناب است، زیرا نشانگر درجه نسبتاً بالای «انسجام آوایی» در درک گفتار است. سیستم درکی، ذره‌های آکوستیکی را که به‌طور معمول با یکدیگر ترکیب نمی‌شوند، به هم می‌چسباند.

پدیده دیگری نیز انسجام آوایی در درک گفتار را نشان می‌دهد. فرض کنید ویدیویی از فردی که "ba"، "da" و "ga" را تولید می‌کند، تهیه می‌کنید. سپس صدای هریک از این هجاها را بر روی تصویر هجای دیگر می‌گذارید. یعنی، یک کپی از ویدیوی [ba] را با صدای [da] هم‌زمان می‌کنید، کپی دیگری از آن را با صدای [ga] و به‌همین ترتیب الی آخر. ابهام‌های جالبی در میان این‌گونه نمونه‌های نامنتطبق صوتی-تصویری دیده می‌شود و یکی از آنها به‌طور ویژه تبدیل به مثال مشهور و جالب انسجام آوایی در درک گفتار شده است.

برخی از این نمونه‌های نامنتطبق صوتی-تصویری هرگز درست به نظر نمی‌رسند. به عنوان مثال، هنگامی که صدای [da] را روی تصویر [ba] می‌گذارید، شنوندگان نمونه "ba" را گزارش می‌دهند (بر اساس حرکت مشخص انسداد لبی) اما این طبیعی به نظر نمی‌رسد.

نمونه نامنتطبق صوتی-تصویری واقعاً جالب، نمونه‌ای است که با گذاشتن صدای [ba] بر روی تصویر [ga] پدید می‌آید. فیلم حاصل به نظر شبیه به هیچ‌یک از دو هجای ورودی نیست، بلکه در عوض شبیه به "da" است! این توهم درکی^۱ اثر مک‌گرک^۲ نامیده می‌شود، بر اساس نام «هری مک‌گرک»^۳ که برای اولین بار آن را نشان داد (مک‌گرک و مک‌دونالد^۴ ۱۹۷۶). این توهم، به طرز شگفت‌آوری قوی است و تنها وقتی از بین می‌رود که چشم‌هایتان را می‌بندید. حتی اگر بدانید که سیگنال صوتی [ba] است، فقط "da" می‌شنوید.

1. perceptual illusion

2. McGurk effect

3. Harry McGurk

4. MacDonald

اثر مک‌گرک مثالی است که نشان می‌دهد چگونه ما در فرایند درک گفتار، از دانش آوایی مان برای ایجاد درکی به‌لحاظ آوایی منسجم استفاده می‌کنیم. ما به‌عنوان شنوندگان اطلاعات گوش‌ها و چشم‌هایمان را ترکیب می‌کنیم تا به قضاوت آوایی درباره آنچه گفته می‌شود، دست پیدا کنیم. این فرایند، دانش آوایی خاصی را به کار می‌گیرد و نه منحصرراً دانش عام حرکات گفتاری را. به‌عنوان نمونه، والکر و همکاران^۱ (۱۹۹۵) (با گذاشتن صدای یک فرد روی چهره فردی دیگر) نشان دادند که تلفیق صدا-تصویر وقتی شنونده، گوینده را می‌شناسند و می‌دانند که صدا به تصویر تعلق ندارد، متوقف می‌شود. این واقعیت نشان می‌دهد که انسجام آوایی یک ویژگی درک گفتار است و اینکه انسجام آوایی یک قابلیت درکی آموختنی است مبتنی بر دانشی که ما به‌عنوان شنوندگان کسب می‌کنیم.

مک‌گرک‌کردن^۲ تا حد افراط

اثر مک‌گرک پدیده‌ای بسیار رایج در درک گفتار است و پژوهشگران این اثر را بسیار مورد بررسی قرار داده‌اند. در واقع، این پدیده به‌قدری رایج است که می‌توان از اسم «اثر مک‌گرک» فعل آن را نیز ساخت - مک‌گرک‌کردن یعنی اثر مک‌گرک داشتن. مثال‌های زیر برخی از موارد مک‌گرک‌کردن را نشان می‌دهد.

نوزادان مک‌گرک می‌کنند (روزن‌بلوم و همکاران ۱۹۹۷)

شما می‌توانید مک‌گرک کنید حتی وقتی تلویزیون سروته است (کمپبل^۳ ۱۹۹۴)

شنوندگان ژاپنی کمتر از شنوندگان انگلیسی مک‌گرک می‌کنند (سیکی‌یاما و توکورا^۴ ۱۹۹۳)

چهره مردان می‌تواند با صدای زنان مک‌گرک شود (گرین و همکاران ۱۹۹۱)

یک چهره آشنا با صدایی ناآشنا مک‌گرک نمی‌شود (والکر و همکاران ۱۹۹۵)

1. Walker et al.

۲. McGurking. مک‌گرک‌کردن یعنی ادغام کردن اطلاعات آوایی با اطلاعات تصویری (مترجم)

3. Campbell

4. Sekiyama and Tohkura

۵-۳. دانش زبانی درک گفتار را شکل می‌دهد

تاکنون دیدیم که بخشی از توانایی ما در درک گفتار به وسیله روابط غیرخطی و سایر ویژگی‌های سیستم شنوایی انسان شکل می‌گیرد و نیز دیدیم هنگام گوش دادن به گفتار آنچه می‌شنویم تا حدی به وسیله دانش آوایی که به عنوان گویشور کسب کرده‌ایم، شکل می‌گیرد. اکنون به احتمال دیگری می‌پردازیم که بر اساس آن درک گفتار همچنین به وسیله دانش ما از ساختار زبان‌شناختی زبان مادری‌مان شکل می‌گیرد.

در بخش ۵-۲ (دانش آوایی) به این واقعیت پرداختیم که مجموعه صداهای گفتاری زبان مادری، درک گفتار را شکل می‌دهد، از این رو، در این بخش وقتی صحبت از «ساختار زبانی» می‌کنم، به دانش واجی نخواهم پرداخت، بلکه در عوض، شواهدی از تأثیر واژگان بر درک گفتار ارائه خواهم کرد. به عبارت دیگر، شنیدن واژه‌ها متفاوت است از شنیدن صداهای گفتاری.

در ابتدا باید بگویم که در این مورد اختلاف نظر وجود دارد. به نظر من درک گفتار تحت تأثیر وضعیت واژگانی الگوهای صوتی که می‌شنویم قرار دارد، اما برخی از زبان‌شناسان با این دیدگاه مخالف‌اند.

روش علمی: در مسیر متقاعدشدن

برای یک استدلال علمی محکم، عناصر بسیاری لازم است که در اینجا به آنها نمی‌پردازم. اما می‌خواهم به نکته‌ای درباره اینکه چگونه پیشرفت می‌کنیم، اشاره کنم. نکته این است که هیچ فردی به تنهایی نمی‌تواند تعیین کند که یک استدلال پیروز شد یا شکست خورد. من وقتی که یک مقاله پژوهشی می‌نویسم، معمولاً تحت تأثیر استدلال‌ها و ذکاوت خود قرار می‌گیرم. فکر می‌کنم چیزی فهمیده‌ام و می‌خواهم نتیجه آن را به همه دنیا اعلام کنم. اما نتیجه واقعی کار من همیشه توسط خوانندگان من نوشته می‌شود و با هر فرد جدیدی که مقاله را می‌خواند، ادامه پیدا می‌کند. اینکه آیا نتیجه قابل قبول یا معتبر است به تصمیم آنها بستگی دارد. این جنبه از روش علمی، به همراه داوری تخصصی مقالات ارسالی برای چاپ، بخشی از آن مسیری است که ما را به سوی پاسخ‌های درست هدایت می‌کند.

این سؤال که آیا درک گفتار تحت تأثیر پردازش کلمه قرار دارد یا نه، از این لحاظ جالب است. پژوهشگران بسیار زبده، بسیار باهوش و توانا در زبان‌شناسی، در این زمینه اختلاف نظر دارند. برخی با یک استدلال یا مجموعه‌ای از نتایج متقاعد می‌شوند و برخی تحت تأثیر مجموعه دیگری از یافته‌ها و یا روش‌های فکری دیگری قرار دارند. آنچه برای من جالب است، این است که این مسئله به‌مدت بسیار بسیار طولانی به درازا کشیده است و از آن جالب‌تر اینکه همان‌طور که مسئله ادامه پیدا می‌کند و پژوهشگران داده‌های بیشتر و بیشتری در رابطه با سؤال بر روی هم انباشته می‌کنند، نظریه‌ها اندک‌اندک شروع به درهم آمیختن می‌کنند. که البته در اینجا به آن نمی‌پردازیم.

شیوه عمل «لغزش‌های گوش» نشان می‌دهد که شنوندگان، دانش خود درباره کلمات را برای درک گفتار به کار می‌گیرند. زینی باند^۱ (۱۹۹۹) به خطاهای درکی مثل “spun toffee” که به‌صورت “fun stocking” شنیده شده و نیز “wrapping service” که به‌صورت “wrecking service” شنیده شده، اشاره می‌کند. در پیکره‌ای که وی از لغزش‌های گوش گردآوری می‌کند، تقریباً همه لغزش‌ها، اشتباهات درکی^۲ کلمه هستند و نه اشتباهات درکی واج. البته گاهی ممکن است یک صدای گفتاری را بد بشنویم و شاید فکر کنیم که گوینده کلمه را اشتباهی تلفظ کرده است، اما پژوهش زینی باند نشان می‌دهد که شنوندگان بی‌وقفه به شنیدن کلمات واداشته می‌شوند، حتی وقتی که فرایند ارتباط مختل می‌شود. این بسیار معنادار است، زیرا هدف ما در ارتباط گفتاری درک گفتار فرد مقابل است و کلمات (یا به‌طور تخصصی‌تر، تکواژها) واحدهایی هستند که وقتی حرف می‌زنیم با یکدیگر ردوبدل می‌کنیم.

این شم زبانی که افراد گرایش به شنیدن کلمات دارند، با ادامه هوشمندانه آزمایش محل تولید که در بخش ۵-۱ و ۵-۲ مورد بحث قرار گرفت، اثبات شد. اثر گانونگ^۳ که بر اساس نام پژوهشگری که برای اولین بار آن را کشف کرد (گانونگ ۱۹۸۰) نام‌گذاری شد، شامل پیوستاری مثل پیوستار شکل ۵-۱ است، با این تفاوت

1. Zinny Bond

2. misperceptions

3. Ganong effect

که یک کلمه در یک انتها و یک ناکلمه در انتهای دیگر آن وجود دارد. به‌طور مثال، اگر به انتهای پیوستار [da]-[ga] یک [g] بیافزاییم، در این صورت پیوستاری بین کلمه «dog» و ناکلمه [gag] خواهیم داشت. آنچه «گانونگ» دریافت و آنچه مرا به این فکر وامی‌دارد که بخشی از درک گفتار به‌وسیله دانش واژگانی شکل می‌گیرد، این است که در این پیوستار جدید، پاسخ‌های «dog» بیشتری در مقایسه با پاسخ‌های «da» در پیوستار [da]-[ga] به دست می‌آوریم. ایده «آهن‌ربای درکی» را به خاطر دارید؟ در اثر «گانونگ»، کلمات مانند آهن‌رباهای درکی عمل می‌کنند؛ بدین‌صورت که هرگاه در یک انتهای پیوستار یک کلمه وجود دارد، شنوندگان گرایش به شنیدن نمونه‌های بیشتری به‌صورت یک واحد واژگانی و نمونه‌های کمتری به‌صورت ناکلمه در انتهای دیگر پیوستار دارند.

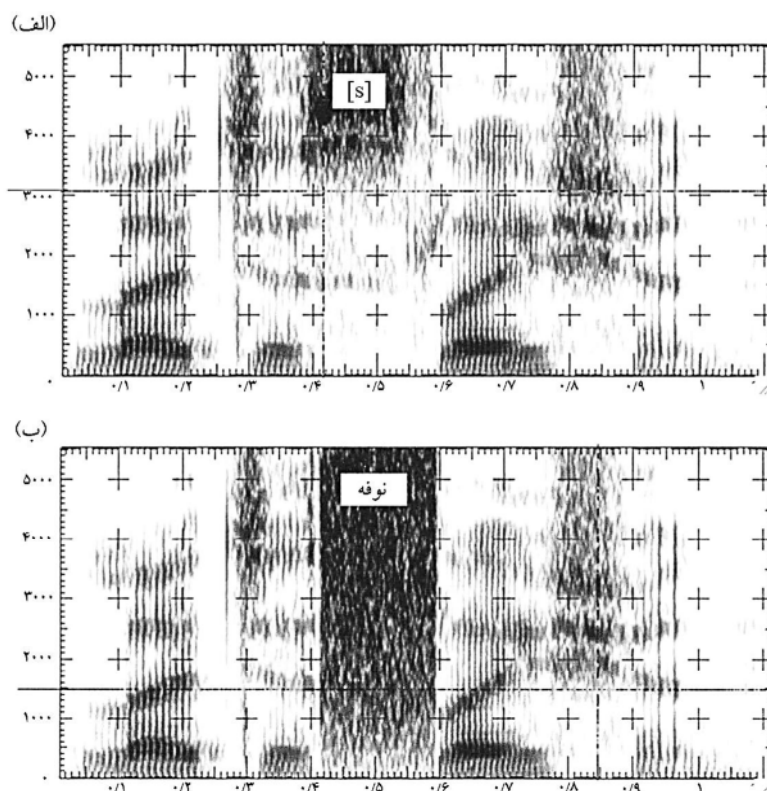
«گانونگ» کنترل آزمایشی دقیقی با استفاده از جفت پیوستارهایی مانند «tash»-«dash» و «task»-«dask» را به کار بست. در این پیوستارها شباهت بسیار زیادی بین پیوستاری که دارای یک کلمه در انتهای /t/ است («task»-«dask») با پیوستاری که دارای یک کلمه در انتهای /d/ است («tash»-«dash»)، مشاهده می‌شود. بنابراین، کمتر محتمل است که اختلاف در تعداد پاسخ‌های «d» مربوط به تفاوت‌های آکوستیکی جزئی بین پیوستارها باشد تا مربوط به واژه‌ای بودن انتهای پیوستارها. همچنین، مشاهده شد که تأثیر واژگانی وقتی که صداها مورد آزمایش در پایان کلمات آزمایش قرار دارند، قوی‌تر است، مثل «kiss»-«kish» در مقابل «fiss»-«fish». این نتیجه منطقی به نظر می‌رسد، زیرا فعال‌شدن یک کلمه در واژگان ذهنی مستلزم اندکی زمان است.

سومین پدیده درکی که نشان می‌دهد دانش زبانی (با ماهیت واژگانی) درک گفتار را شکل می‌دهد، توسط وارن^۱، کاشف آن، «بازگردانی واج»^۲ نامیده شد (وارن ۱۹۷۰). شکل ۵-۷ بازگردانی واج را نشان می‌دهد. در بالای این شکل، طیف‌نگاشت کلمه «legislation» را می‌بینید. شکل پایین نیز طیف‌نگاشت همین کلمه است با این تفاوت که در آن نوفه‌ای با پهنای باند وسیع جایگزین [s] شده است. هنگامی که افراد

1. Warren

2. phoneme restoration

به سیگنالی که با نوفه جایگزین شده است (شکل ۵-۷ (ب))، گوش می‌دهند، [s] در [ˌleɪdʒɪsˈleɪʃən] را «می‌شنوند». آرتور ساموئل^۱ (۱۹۹۱) در مطالعه خود اظهار می‌کند که آوای [s] واقعاً در نمونه‌های جایگزین شده با نوفه درک می‌شود. او دریافت که شنوندگان واقعاً قادر نیستند تفاوت بین سیگنالی که نوفه به آن افزوده شده (نوفه‌ای با پهنای باند وسیع به [s] موجود افزوده شده است) با سیگنالی که نوفه در آن جایگزین شده (ابتدا [s] حذف شده سپس نوفه افزوده می‌شود) را تشخیص دهند. این بدان معناست که [s] در واقع درک یا بازگردانی می‌شود و بنابراین، دانش شما از کلمه “legislation” درک شما از این نوفه رهش را شکل داده است.



شکل ۵-۷. (الف) طیف‌نگاشت کلمه “legislation” که در آن نوفه [s] مشخص شده است. (ب) طیف‌نگاشت همان کلمه با این تفاوت که [s] با نوفه‌ای با پهنای باند وسیع جایگزین شده است.

جف المان^۱ و جی مک‌کلیند^۲ (۱۹۸۸) شاهد مهم دیگری مبنی بر اینکه دانش زبانی درک گفتار را شکل می‌دهد، ارائه می‌کنند. در آزمایش آنها فرایند بازگردانی واج، منجر به درک صدایی می‌شود که بعداً در یک فرایند جبران هم‌تولیدی مشارکت می‌کند. این فرایند دومرحله‌ای اندکی پیچیده است، اما یکی از هوشمندانه‌ترین و تأثیرگذارترین آزمایش‌ها در این زمینه است.

مرحله اول: جبران هم‌تولیدی. پیوستاری از [da]-[ga] دقیقاً مانند پیوستار شکل ۵-۱ می‌سازیم با این تفاوت که به‌جای بافت هجایی [al] و [ar] از هجاهای [as] و [qf] استفاده می‌کنیم. جبران هم‌تولیدی در بافت هجاهای سایشی نیز روی می‌دهد و همان تأثیری را دارد که در بافت همخوان‌های روان دارد. شنوندگان هجاهای «ga» بیشتری در بافت [as] نسبت به وقتی که بافت [qf] است، می‌شنوند.

مرحله دوم: بازگردانی واج. نوفه‌های سایشی کلمات “abolish” و “progress” را با نوفه‌ای با پهنای باند وسیع جایگزین می‌کنیم، همان‌طور که در مورد [s] در کلمه “legislation” در شکل ۵-۷ جایگزین کردیم. اما درک این کلمات به این صورت است که در “progress” یک [s] درک می‌شود و در “abolish” یک [f]، در حالی که در انتهای موج صوتی این کلمات فقط نوفه وجود دارد.

سؤال این است که آیا بازگردانی [s] و [f] در “progress” و “abolish” حقیقتاً یک پدیده درکی است یا چیزی بیشتر شبیه به سوگیری شنوندگان هنگام تصمیم‌گیری که نشان می‌دهد چگونه آنها ماهیت یک کلمه را حدس می‌زنند. آیا وجود کلمه‌ای مثل “progress” و نبود کلمه‌ای مثل “progresh” واقعاً بر درک گفتار تأثیر می‌گذارد؟ المان و مک‌کلیند برای بررسی این سؤال، در یک آزمایش، “abolish” و “progress” را به‌عنوان بافت آزمایش جبران هم‌تولیدی به کار بردند. استدلال آنها این بود که اگر [s] «بازگردانی‌شده» اثر جبران هم‌تولیدی ایجاد کند، به‌نحوی که وقتی در بافت ماقبل یک [s] بازگردانی‌شده وجود داشته باشد، شنوندگان هجاهای “ga” بیشتری نسبت به وقتی

1. Jeff Elman

2. Jay McClelland

که یک [ʃ] بازگردانی شده باشد، بشنوند. در این صورت، باید چنین نتیجه گرفت که شنوندگان واقعاً [s] و [ʃ] را درک می‌کنند و در واقع، این آواها به لحاظ درکی در آنجا حضور دارند و قادرند بر درک پیوستار [da]-[ga] تأثیر متقابل بگذارند. حدس بزنید المان و مک‌کلیند چه نتایجی به دست آوردند؟ [s] و [ʃ] خیالی که در واقع در آنجا نبودند، سبب جبران هم‌تولیدی شدند، شاهد بسیار قابل توجهی از اینکه درک گفتار به وسیله دانش زبانی شکل می‌گیرد.

۵-۴. شباهت درکی^۱

اکنون در پایان فصل، مایلم روشی برای اندازه‌گیری فضاهای شباهت درکی آواهای گفتاری مطرح کنم. این روش در فصل‌های آینده که انواع مختلف آواها، ویژگی‌های آکوستیکی و سپس شباهت‌های درکی آنها مورد بحث قرار می‌گیرد، مفید خواهد بود. شباهت درکی همچنین، عاملی کلیدی است در ربط ویژگی‌های آوایی به تغییرات صوتی زبان و نیز در الگوهای واجی زبان که ناشی از تغییرات صوتی هستند. در این روش، هجاهای آزمایش در اختیار شنوندگان قرار می‌گیرد و از آنها خواسته می‌شود که صداها را در هجاها شناسایی کنند. معمولاً به دلیل تولید دقیق «گفتار آزمایشگاهی» (گفتاری که با خواندن فهرستی از هجاها با یک میکروفن در آزمایشگاه آواشناسی تولید می‌شود)، شنوندگان اشتباهات شناسایی بسیار کمی در این وظیفه مرتکب می‌شوند. از این رو، ما معمولاً اندکی نوفه به هجاهای آزمایش برای ایجاد مقداری خطا اضافه می‌کنیم. سطح نوفه به صورت نسبت شدت نوفه در مقایسه با شدت قله هجا اندازه‌گیری می‌شود که به آن نسبت سیگنال به نوفه می‌گویند و برحسب دسی بل است. برای تحلیل پاسخ شنوندگان، آنها را در یک ماتریس ابهام^۲ فهرست می‌کنیم. در هر ردیف ماتریس، یکی از هجاهای آزمایش (سرجمع هر ۱۰ نمونه آن هجا) فهرست می‌شود و در هر ستون ماتریس، یکی از پاسخ‌هایی که در اختیار شنوندگان است.

1. perceptual similarity

2. confusion matrix

جدول ۵-۲. ابهام سایشی‌ها (و [d]) در آزمایش میلر و نایسلی (۱۹۵۵).

کل	سایر	“d”	“z”	“s”	“dh”	“th”	“v”	“f”
۲۶۴	۱۴	۰	۰	۴	۱	۴۶	۰	۱۹۹
۲۳۶	۲۲	۰	۴	۰	۲۹	۱	۱۷۷	۳
۲۳۲	۲۱	۰	۰	۱۰	۰	۱۱۴	۲	۸۵
۲۰۴	۱۷	۰	۱۸	۰	۱۰۵	۰	۶۴	۰
۲۲۸	۱۵	۰	۰	۱۷۰	۰	۳۸	۰	۵
۲۲۴	۴۹	۱۷	۱۳۲	۰	۲۲	۰	۴	۰
۲۶۰	۵۹	۱۸۹	۸	۰	۴	۰	۰	۰

جدول ۵-۲، ماتریس ابهام حاصل از تحقیق گسترده جرج میلر^۱ و پاتریشیا نایسلی^۲ (۱۹۵۵) در زمینه درک همخوان با نسبت سیگنال به نوفه صفر دسی‌بل را نشان می‌دهد. این داده‌ها هرچند قدیمی، اما خوب هستند. با نگاه به اولین ردیف ماتریس ابهام، می‌بینیم که [f] ۲۶۴ بار تولید شده است: ۱۹۹ بار به‌درستی به صورت “f” و ۴۶ بار به غلط به‌صورت “th” شناسایی شده است. لازم به ذکر است که «میلر و نایسلی» برای برخی صداها داده‌های بیشتری در مقایسه با برخی دیگر دارند. بر اساس ماتریس ابهام می‌توان پاسخ نسبتاً سریعی برای برخی سؤال‌ها یافت، حتی قبل از آنکه هر تجزیه و تحلیل دقیقی بر روی داده‌ها صورت گیرد. به‌عنوان مثال، چرا کودکان گاهی اسم “Keith” را به‌صورت “Keif” تلفظ می‌کنند؟ بر اساس داده‌های «میلر و نایسلی»، [θ] ۸۵ بار از ۲۳۲ بار، به‌صورت “f” شنیده شده، ابهام با “f” بیش از هر صدای گفتاری دیگری در آزمایش رخ داده است. اما روشن نیست که آیا این داده‌ها اصلاً چیزی درباره سایر نکات مورد توجه ما می‌گویند یا نه. به‌عنوان مثال، چرا “this” و “that” گاهی اوقات با صدای [d] تلفظ می‌شوند؟ برای پاسخ به این سؤال باید روشی برای ترسیم نقشه «فضای» درکی که زیربنای ابهامات ملاحظه‌شده در آزمایش است، پیدا کنیم. اکنون به رسم نقشه می‌پردازیم.

1. George Miller

2. Patricia Nicely

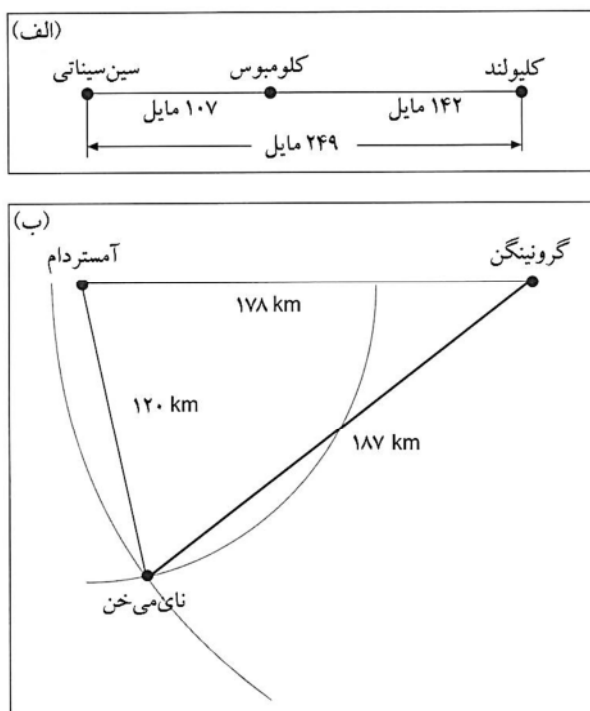
۵-۴-۱. رسم نقشه بر اساس فاصله

با استخراج اطلاعات از یک ماتریس ابهام، سعی می‌کنیم تصویری از سیستم درکی که سبب بروز ابهام می‌شود، به دست آوریم. در روشی که به کار می‌بریم، ابتدا فهرستی از فاصله‌ها را انتخاب کرده، سپس آنها را به صورت یک نقشه بازسازی می‌کنیم. به عنوان مثال، فهرست فاصله‌های زیر را برای شهرهای اوهایو در نظر بگیرید.

کلمبوس^۱ تا سین سیناتی^۲، ۱۰۷ مایل

کلمبوس تا کیولند^۳، ۱۴۲ مایل

سین سیناتی تا کیولند، ۲۴۹ مایل



شکل ۵-۸ (الف) نقشه یک بعدی سه شهر در اوهایو. یک بُعد به خوبی فاصله بین آنها را نشان می‌دهد. (ب) نقشه دوبعدی سه شهر در هلند. قوس‌ها نشان می‌دهند که چطور جای نای می‌خن بر روی نقشه تعیین شد.

با استفاده از این فاصله‌ها و با گذاشتن «کلمبوس» بین «کلیولند» و «سین سیناتی» می‌توانیم شهرهای مذکور را به روی یک خط مستقیم مثل شکل ۵-۸ (الف) رسم کنیم. یک خط برای نشان دادن این فاصله‌ها کافی است، زیرا فاصله بین «سین سیناتی» تا «کلیولند» برابر با مجموع دو فاصله دیگر است ($249 = 142 + 107$).

اما برای رسم مثال زیر به یک سطح دوبعدی نیاز است.

آمستردام به گرونینگن^۱، ۱۷۸ کیلومتر

آمستردام به نای‌می‌خن^۲، ۱۲۰ کیلومتر

گرونینگن به نای‌می‌خن، ۱۸۷ کیلومتر

شکل ۵-۸ (ب) نقشه دوبعدی فاصله بین این شهرهای هلند را نشان می‌دهد. برای رسم این نقشه، «آمستردام» و «گرونینگن» را روی یک خط با فاصله ۱۷۸ کیلومتر رسم کردم. سپس، یک قوس با فاصله ۱۲۰ کیلومتر از آمستردام رسم کردم، با آگاهی از اینکه نای‌می‌خن باید جایی روی این قوس باشد. یک قوس دیگر نیز با فاصله ۱۸۷ کیلومتر از گرونینگن رسم کردم، با آگاهی از اینکه نای‌می‌خن باید جایی روی این قوس است. بنابراین، «نای‌می‌خن» باید با فاصله ۱۲۰ کیلومتر از «آمستردام» و ۱۸۷ کیلومتر از «گرونینگن» در تقاطع این دو قوس باشد. به این روش، یعنی تعیین مکان یک نقطه سوم بر حسب فاصله آن از دو نقطه معین، مثلث‌بندی^۳ می‌گویند. مثلث شکل ۵-۸ (ب)، همان‌طور که در نقشه شکل ۵-۹ نیز می‌توان دید، تصویر دقیقی از موقعیت نسبی این سه شهر را نشان می‌دهد.

ممکن است با خود فکر کنید: «خوب، همه اینها بسیار عالی، اما چه ارتباطی با درک گفتار دارد؟» سؤال خوبی است. به نظر می‌رسد که می‌توانیم فاصله‌های درکی را از یک ماتریس ابهام به دست آوریم و سپس، با استفاده از روشی در ادامه مثلث‌بندی به نام مقیاس‌گذاری چندبعدی^۴ می‌توانیم یک نقشه درکی از ماتریس ابهام به دست آوریم.

1. Groningen

2. Nijmegen

3. triangulation

4. multidimensional scaling (MDS)



شکل ۵-۹. نقشه هلند که موقعیت مثلث آمستردام-گرونینگن-نای می‌خن در شکل ۵-۸ (ب) را نشان می‌دهد.

۵-۴-۲. نقشه درکی همخوان‌های سایشی

در این بخش برای رسم نقشه فضای درکی که سبب ایجاد الگوی ابهام در جدول ۵-۲ شده، از روش مقیاس‌گذاری چندبعدی (ام‌دی‌اس) استفاده خواهیم کرد. اولین مرحله این فرایند، تبدیل ابهام‌ها به فاصله‌هاست که به نظر من عملی منطقی است، زیرا چنین فرض می‌کنیم که وقتی آواها در فضای درکی به یکدیگر نزدیک هستند، در وظیفه شناسایی^۱ با یکدیگر اشتباه می‌شوند. بنابراین، خطاهای

1. identification task

ماتریس ابهام در جدول ۵-۲ به ما می‌گوید که چه آوایی با چه آوای دیگری اشتباه شده است. برای مثال، ابهام همخوان‌های واک‌دار [v]، [ð]، [z] و [d] خیلی به‌ندرت با همخوان‌های بی‌واک [f]، [θ] و [s] روی می‌دهد. این نشان می‌دهد که همخوان‌های واک‌دار در فضای درکی نزدیک به یکدیگر هستند و همخوان‌های بی‌واک نواحی دیگری را اشغال می‌کنند. نتایج کلی از این دست، همگی خوب هستند، اما باید برآورد مشخصی از فاصله درکی بر اساس ماتریس ابهام به دست آوریم.

برای این کار می‌توان روش زیر را به کار برد (من روش روان‌شناس ریاضیدان، «روجر شپارد»^۱ را که در مقاله سال ۱۹۷۲ وی با عنوان «بازنمایی روان‌شناختی صداهای گفتاری» مطرح شد، به کار می‌برم). این روش شامل دو مرحله است: ابتدا شباهت را محاسبه کنید و سپس فاصله‌ها را از شباهت‌ها استخراج کنید.

محاسبه شباهت آسان است. تعداد دفعاتی که فکر می‌کنید [f] شبیه به «θ» شنیده می‌شود، منعکس‌کننده شباهت «f» و «θ» در فضای درکی شماست. علاوه بر آن، شباهت «θ» - «f» در تعداد دفعاتی که [θ] شبیه به «f» شنیده می‌شود نیز انعکاس پیدا می‌کند. از این رو، این دو سلول را در ماتریس ابهام ترکیب می‌کنیم: [f] به‌صورت «θ» شنیده می‌شود و «θ» به‌صورت «f» شنیده می‌شود. از آنجا که ممکن است تعداد نمونه‌های [f] و [θ] در داده‌ها متفاوت باشد، ما نسبت آنها را به کار می‌بریم نه تعداد مطلق آنها را.

دقت کنید که برای هر دو نمونه در ماتریس ابهام، زیرماتریسی با چهار سلول داریم: زیرماتریس (الف) نسبت پاسخ‌های تقابل «θ» «f» در داده‌های میلر و نایسلی را نشان می‌دهد. به‌عنوان مثال، مقدار ۰/۷۵ در این جدول نسبت نمونه‌های [f] است که به‌صورت «f» شناسایی شده‌اند (۰/۷۵۴ = ۱۹۹/۲۶۴). همراه با زیرماتریس (الف)، دو صورت انتزاعی‌شده آن نیز آمده است.

(الف)		(ب)		(ج)			
«f»	«θ»	«f»	«θ»	«i»	«j»		
[f]	۰/۷۵	[f]	p _{ff}	[i]	p _{ii}	p _{ij}	
[θ]	۰/۳۷	[θ]	p _{θf}	[j]	p _{ji}	p _{jθ}	

1. Roger Shepard

متغیرهای زیرماتریس (ب) نسبت‌ها را به صورت کد نشان می‌دهند؛ بدین صورت که “p” یعنی نسبت، اولین اندیس نشانهٔ برجسب ردیف و دومین اندیس نشانهٔ برجسب ستون است. از این رو، $p_{\theta f}$ متغیری است که به نسبت دفعاتی اشاره دارد که نمونه‌های $[\theta]$ به صورت “f” نامیده شده‌اند. مقدار $p_{\theta f}$ در این داده‌ها برابر با ۰/۳۷ است. زیرماتریس (ج) از این هم اندکی انتزاعی‌تر است. بر اساس آن برای هر دو صدای i و j، زیرماتریسی با ابهام‌ها (اندیس‌هایی که مثل هم نیستند) و پاسخ‌های درست (اندیس‌هایی که مثل هم هستند) داریم.

عدم تقارن در ماتریس‌های ابهام

آیا در اینکه $[\theta]$ خیلی بیشتر به صورت “f” شنیده شده تا $[f]$ به صورت “th”، معنای بسیار مهمی وجود دارد؟ ممکن است شنوندگان نوعی سوگیری در برابر نامیدن صداها به صورت “th” داشته‌اند، شاید به این دلیل که تمایز بین “th” و “dh” بر روی ورقهٔ پاسخ گیج‌کننده بوده است. به نظر می‌رسد که در جدول ۵-۲ همین مسئله روی داده باشد، زیرا پاسخ‌های “f” در مجموع در مقایسه با پاسخ‌های “th” خیلی بیشترند. اما کم بودن نسبی پاسخ‌های “s” نشان می‌دهد که نمی‌توان خیلی روی توجیه سوگیری تکیه کرد، زیرا نگاشت “s” به $[s]$ در انگلیسی رایج و بدون ابهام است. یک نکتهٔ جالب دربارهٔ عدم تقارن ابهام‌های $[f]$ و $[\theta]$ این است که این ابهام درکی با گرایش جهانی تحول آوایی هماهنگی دارد ($[\theta]$ بیشتر به $[f]$ تحول پیدا می‌کند تا خلاف آن). آیا این فقط یک تصادف است یا یک رابطهٔ علی وجود دارد؟ روش شپارد برای محاسبهٔ شباهت از یک ماتریس ابهام بر این نکتهٔ جالب سرپوش می‌گذارد و چنین فرض می‌کند که $p_{\theta f}$ و $p_{f\theta}$ دو مقدار ناقص یک چیز، میزان ابهام‌زایی^۱ “f” و “θ” هستند. این دو مقدار سپس با یکدیگر ترکیب می‌شوند تا از میزان شباهت “f”-“θ” یک مقدار به دست دهند. قابل انکار نیست که ممکن است نکتهٔ جالب توجهی در عدم تقارن وجود داشته باشد، اما باید بگویم از آنجا که هدف، رسم نقشه‌های درکی است، ریشه‌های عدم تقارن در ماتریس ابهام نادیده گرفته شده است.

روش شپارد برای محاسبه شباهت از یک ماتریس ابهام به صورت زیر است. ابهام‌های بین دو صدا را به دست می‌آوریم. سپس، نسبت آنها را به پاسخ‌های صحیح حساب می‌کنیم. در ریاضی یعنی:

$$S_{ij} = \frac{P_{ij} + P_{ji}}{P_{ii} + P_{jj}}$$

در این فرمول، مقدار S_{ij} برابر با میزان شباهت بین مقوله i و مقوله j است. در مورد نمونه‌های "P" و "θ" در داده‌های میلر و نایسلی (جدول ۵-۲) محاسبه شباهت به صورت زیر است:

$$S_{ij} = 0/43 = \frac{0/17 + 0/37}{0/75 + 0/49}$$

شپارد درباره این فرمول خیلی کوتاه می‌گوید «فرمول کاربردی است». گاهی اوقات می‌توان همین نتایج را صرفاً با محاسبه میانگین دو نسبت P_{ij} و P_{ji} به عنوان اندازه شباهت به دست آورد. اما فرمول شپارد در ماتریس ابهامی عملکرد بهتری دارد که در آن یک مقوله، ابهام‌هایی دارد که فقط بین دو پاسخ خاص متمرکز شده‌اند، در حالی که مقوله دیگر ابهام‌هایی دارد که تقریباً به طور گسترده‌ای بین پاسخ‌های ممکن توزیع شده‌اند؛ مثلاً وقتی در مقابل کاربرد یک پاسخ خاص، سوگیری وجود دارد. روش به دست آوردن مقدار شباهت از یک ماتریس ابهام شرح داده شد. برای به دست آوردن فاصله درکی از شباهت، مقدار منفی لگاریتم طبیعی^۱ شباهت را حساب کنید:

$$d_{ij} = -\ln(S_{ij})$$

این فرمول مبتنی بر قانون شپارد است. بر اساس این قانون، رابطه بین فاصله درکی و شباهت، نمایی^۲ است. ممکن است حقیقتی ژرف در باب پردازش ذهن در این قانون وجود داشته باشد، این قانون در مباحث نامرتبط مختلفی مطرح می‌شود (شانون و ویور^۳ ۱۹۴۹، پارزن^۴ ۱۹۶۲) که البته این موضوع دیگری است.

1. natural log

2. exponential

3. Shannon and Weaver

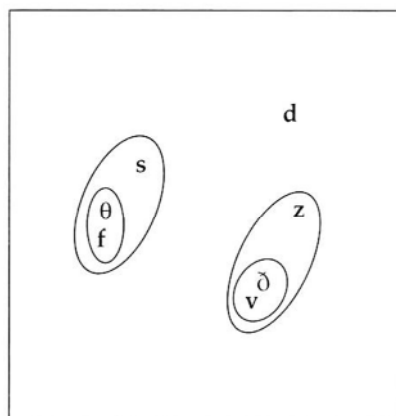
4. Parzen

جدول ۳-۵. شباهت بین سایشی‌های انگلیسی آمریکایی (و [d]) بر مبنای ماتریس ابهام میلر و نایسلی (۱۹۵۵) که نسبت سیگنال به نوفه آن صفر دسیبل است.

	“f”	“v”	“th”	“dh”	“s”	“z”	“d”
[f]	۱/۰						
[v]	۰/۰۰۸	۱/۰					
[θ]	۰/۴۳۴	۰/۰۱۰	۱/۰				
[ð]	۰/۰۰۳	۰/۳۴۵	۰/۰۰۰	۱/۰			
[s]	۰/۰۲۵	۰/۰۰۰	۰/۱۷۰	۰/۰۰۰	۱/۰		
[z]	۰/۰۰۰	۰/۰۲۶	۰/۰۰۰	۰/۱۶۹	۰/۰۰۰	۱/۰	
[d]	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۱۲	۰/۰۰۰	۰/۰۸۱	۱/۰

به هر روی، اکنون به موضوع رسم نقشه بر می‌گردیم با این تفاوت که به‌جای رسم موقعیت شهرهای هلند در فضای جغرافیایی، نقشه فضای درکی همخوان‌های سایشی و “d” را در زبان انگلیسی رسم می‌کنیم. جدول ۳-۵ مقدار شباهت‌های به‌دست‌آمده از ماتریس ابهام «میلر و نایسلی» (جدول ۲-۵) را با استفاده از معادله (۱-۵) نشان می‌دهد.

نقشه درکی به‌دست‌آمده بر پایه این شباهت‌ها در شکل ۵-۱۰ رسم شده است. یکی از اولین نکات قابل توجه درباره این نقشه این است که همخوان‌های واک‌دار در یک طرف و همخوان‌های بی‌واک در طرف دیگر آن قرار دارند. این مؤید نظر اولیه ماست که با ملاحظه ابهام‌های خام به آن رسیدیم؛ اینکه صداهای بی‌واک به‌ندرت به‌صورت واک‌دار نامیده شده‌اند و برعکس. نکته جالب دیگر این است که سایشی‌های واک‌دار و بی‌واک به یک ترتیب بر روی محور عمودی قرار گرفته‌اند. این، ممکن است یک بُعد پیشین-پسین باشد یا یک همبستگی جالب با جنبه‌ای آکوستیکی در آواها.



شکل ۵-۱۰. نقشه درکی سایشی ها و [d] در زبان انگلیسی. موقعیت آواها با روش مقیاس گذاری چندبعدی بر اساس داده های جدول ابهام میلر و نایسلی (۱۹۵۵) به دست آمد. گروه های آوایی درون بیضی ها، خوشه هایی هستند که با تحلیل خوشه ای سلسله مراتبی همان داده ها به دست آمدند.

من گرداگرد برخی خوشه های صوتی در شکل ۵-۱۰، بیضی رسم کردم. این بیضی ها دو سطح شباهت در بین آواها را نشان می دهند که با یک روش تحلیل خوشه ای سلسله مراتبی^۱ (روشی دیگر برای تحلیل داده های خام که در بسیاری از نرم افزارهای آماری یافت می شود، برای اطلاعات بیشتر به جانسون ۲۰۰۸ رجوع کنید) به دست آمد. در اولین مرحله خوشه بندی، "θ" با "f" و "v" با "ð" در نقشه درکی تشکیل خوشه می دهند. در سطحی نسبتاً کلی تر، همخوان های صغیری در کنار همخوان های غیرصغیری مجاور خود قرار می گیرند ("s" به خوشه بی واک و "z" به خوشه واک دار می پیوندد). در مرحله بعدی خوشه بندی که در شکل دیده نمی شود، [d] در کنار سایشی های واک دار قرار می گیرد.

ترکیب تحلیل خوشه ای با مقیاس گذاری چندبعدی (ام دی اس)، تصویری تقریباً واضح از نقشه درکی در اختیار ما قرار می دهد. لازم به یادآوری است که این روش ها فقط تکنیک های تصویرسازی داده ها هستند و ما هیچ اطلاعاتی به آنچه که در ماتریس ابهام وجود داشت، نیفزودیم (با این وجود، استفاده از یک فضای دوبعدی برای نشان دادن الگوی ابهام در این صداها طبق تصمیم ما بود).

در رابطه با تولید “this” و “that”، این نتایج نشان می‌دهد که تناوب‌های [d]-[ð] و [z]-[ð] فقط به وسیله شباهت شنیداری-درکی هدایت نمی‌شوند. به طور قطع عوامل دیگری نیز در کارند، زیرا در غیر این صورت می‌بایست “vis” و “vat” به عنوان تولیدهای “this” و “that” مشاهده می‌شدند.

مقیاس گذاری چندبعدی و آواشناسی آکوستیک

یکی از معماهای اصلی در آواشناسی آکوستیک، این است که چه جنبه‌هایی از سیگنال گفتاری مهم هستند و چه جنبه‌هایی اهمیتی ندارند. به یک طیف‌نگاشت نگاه می‌کنید و چیزی در آن می‌بینید. سؤال این است: آیا شنوندگان به آن جزء آوایی اهمیتی می‌دهند یا نه؟ آیا آن جزء اهمیتی دارد؟ آواشناسان به سؤال «آیا آن جزء اهمیتی دارد؟» به روش‌های مختلفی پاسخ داده‌اند.

به عنوان مثال، ما به طیف‌نگاشت‌های بسیاری نگاه می‌کنیم و در رابطه با آن جزء اسرارآمیز می‌پرسیم، «آیا آن جزء همیشه در آنجا هست؟» یکی از حقایق تثبیت شده آواشناسی این است که اگر یک مشخصه آکوستیکی همیشه و یا حتی معمولاً حاضر باشد، شنوندگان منتظر آن در درک هستند. این، حتی در مورد اصطلاحاً پالس‌های بزاقی^۱ که گاهی اوقات در طیف‌نگاشت سایشی‌های کناری [H] و [ʒ] دیده می‌شود، نیز صادق است. (یک پالس بزاقی شبیه به نوفه انفجار رهش یک انسدادی است (فصل ۸ را ببینید)، با این تفاوت که در میان نوفه یک سایشی رخ می‌دهد.) این سایشی‌ها اندکی بزاقی هستند اما به نظر می‌رسد این جنبه نسبتاً فرعی تولید آنها برای درک مفید واقع می‌شود.

پاسخ دیگر به سؤال «آیا آن جزء اهمیتی دارد؟» یافتن منشأ آن جزء در نظریه آکوستیکی تولید گفتار است. به عنوان مثال، گاهی پژواک اتاق می‌تواند سایه‌هایی به یک طیف‌نگاشت «بیفزاید». (در واقع در روزگار ضبط صوت‌های حلقه‌به‌حلقه،

←

۱. spit spike. spike در واقع همان پالس یا تکانه (impulse) است، یک موج گذرا با دیرش کوتاه. (مترجم)

می‌بایست مراقب سایه‌های مغناطیسی باشیم. این سایه‌ها وقتی تصویر مغناطیسی صدا در امتداد لایه‌های نوار بر روی حلقه منتقل می‌شد، ایجاد می‌شدند.) اگر شما نظریه‌ای در زمینه ارتباط بین تولید گفتار و آکوستیک گفتار داشته باشید، می‌توانید به سؤال، به این صورت پاسخ دهید که «آن جزء اهمیتی ندارد، زیرا گوینده آن را تولید نکرده است». در فصل‌های آتی این کتاب، نظریه آکوستیکی تولید گفتار را تا حدی بحث خواهیم کرد.

یکی از پاسخ‌های مورد علاقه من به «آیا آن جزء اهمیتی دارد؟»، «قانون کوپر» است. فرانکلین کوپر^۱ در مقاله ۱۹۵۱ خود با ال لیبرمن^۲ و جان بورست^۳ به نقد مسئله جستجوی «همبسته‌های آکوستیکی»^۴ برای درک گفتار می‌پردازند. آنها اظهار می‌کنند «سؤال‌های بسیاری در زمینه رابطه بین محرک آکوستیکی با درک شنیداری وجود دارد که نمی‌توان صرفاً با بررسی طیف‌نگاشت‌ها به آنها پاسخ داد، قطع نظر از اینکه چقدر طیف‌نگاشت‌ها می‌توانند متعدد و متنوع باشند» (نکته‌ای مهم، قابل توجه فناوران گفتار). در عوض، به گفته آنها «اغلب لازم است تغییرات طیف‌نگاشت کنترل شود و سپس ارزیابی‌ای از تأثیر این تغییرات بر روی صدایی که شنیده می‌شود، صورت گیرد. برای این منظور، ما ابزاری ساختیم» (یکی از اولین سستزکننده‌های گفتار). این یک پاسخ کاملاً صریح است. آیا آن جزء اهمیت دارد؟ خوب، گفتار را بدون آن سنتز کنید و بینید گفتار تغییر می‌کند یا نه.

و سرانجام مقیاس‌گذاری چندبعدی به آن پاسخ می‌دهد. نقشه فضای درکی را رسم می‌کنیم و سپس، همبستگی بین ابعاد نقشه با ویژگی‌های آکوستیکی مورد نظرمان (مثلاً جزء مرموز) را پیدا می‌کنیم. اگر یک مشخصه آکوستیکی همبستگی محکمی با یک بعد درکی دارد، در این صورت می‌توانیم بگوییم که احتمالاً آن مشخصه اهمیت دارد. امتیاز این روش این است که بر اساس گفتار طبیعی است و دیگر اینکه امکان بررسی هم‌زمان پارامترهای آکوستیکی بسیاری را میسر می‌سازد.

1. Franklin Cooper

2. Al Liberman

3. John Borst

4. acoustic correlate

منابعی برای مطالعه بیشتر

- Best, C. T. (1995) A direct realist perspective on cross-language speech perception. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research*, Timonium, MD: York Press, 167-200.
- به توصیف نظریه‌ای درمورد درک زبان گذر گفتار می‌پردازد. براساس این نظریه، شنوندگان آواهای جدید و ناآشنا را بر فهرست آواهای زبان مادری خود می‌نگارند.
- Bond, Z. S. (1999) *Slips of the Ear: Errors in the Perception of Casual Conversation*, San Diego Academic Press.
- گردآوری و تحلیل اشتباهات درکی موجود در مکالمات روزمره فیلم «طبیعت وحشی».
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, MA: MIT Press.
- نظریه و شواهد نظریه گشتالتی شنوایی، کتابی بسیار مهم.
- Campbell, R. (1994) Audiovisual speech: Where, what, when, how? *Current Psychology of Cognition*, 13, 76-80.
- در زمینه انعطاف‌پذیری درکی اثر مک‌گرک.
- Cole, R. A. (1973) Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics*, 13, 153-6.
- مطالعه‌ای که نشان می‌دهد افراد اغلب اشتباهات تلفظی را در ارتباط گفتاری نمی‌شنوند.
- Cooper, F. S., Liberman, A. M., and Borst, J. M. (1951) The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of Science*, 37, 318-25.
- منبع «قانون کوپر».
- Elman, J. L. and McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27, 143-65.
- یکی از هوشمندانه‌ترین و بحث‌برانگیزترین آزمایش‌های درک گفتار که تاکنون گزارش شده است.
- Flege, J. E. (1995) Second language speech learning: Theory, findings, and problems. In w. Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research*, Timonium, MD: York Press, 167-200.
- به توصیف نظریه‌ای درمورد درک زبان گذر گفتار می‌پردازد. براساس این نظریه، شنوندگان آواهای جدید و ناآشنا را بر فهرست آواهای زبان مادری خود می‌نگارند.
- Ganong, W. F. (1980) Phonetic categorization in auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 110-25.
- گزارشی بسیار تأثیرگذار از اینکه چگونه افراد به شنیدن واژه‌ها در درک گفتار واداشته می‌شوند. نتیجه اصلی آن اکنون به عنوان «اثر گانونگ» شناخته می‌شود.
- Green, K. P., Kuhi, P. K., Meltzoff, A. N., and Stevens, E. B. (1991) Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, 50, 524-36.
- تلفیق صداها با چهره‌های نامنتطبق به لحاظ جنسیت در اثر مک‌گرک.

- Jakobson, R., Fant, G., and Halle, M. (1952) *Preliminaries to Speech Analysis*, Cambridge, MA: MIT Press.
- یک کتاب کلاسیک در آواشناسی و واج‌شناسی که در آن مجموعه‌ای از مشخصه‌های تمایزدهنده واجی به لحاظ آکوستیکی توصیف می‌شوند.
- Johnson, K. and Ralston, J. V. (1994) Automaticity in speech perception: Some speech/nonspeech comparisons. *Phonetica*, 51(4), 195-209.
- مجموعه‌ای از آزمایش‌هایی که نشان می‌دهند یادگیری مکرر ناشی از برخی «ویژگی‌های خاص» درک گفتار است.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992) Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science*, 255, 606-8.
- اثر آه‌ربای درکی را در کودکان نشان می‌دهد.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957) The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358 - 68.
- توصیف کلاسیک درک مقوله‌ای در درک گفتار.
- Lotto, A. J. and Kluender, K. R. (1998) General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics*, 60, 602-19.
- توجیه اینکه حداقل بخشی از اثر جبران هم‌تولیدی (مان، ۱۹۸۰) ناشی از تقابل شنیداری است.
- Mann, V. A. (1980) Influence of preceding liquid on stop-consonant perception. *Perception & Psychophysics*, 28, 407-12.
- اولین مقاله‌ای که جبران هم‌تولیدی در توالی‌هایی مثل [al da] و [ar ga] را ارائه کرد.
- McGurk, H. and MacDonald, J. (1976) Hearing lips and seeing voices. *Nature*, 264, 746-8.
- اثر شنیداری-تصویری درک گفتار که در این مقاله گزارش شد، در نهایت اثر مک‌گرگ نامیده شد.
- Miller, G. A. and Nicely, P. E. (1955) An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338-52.
- یک منبع استاندارد برای ابهام‌زایی آواهای گفتاری زبان انگلیسی آمریکایی.
- Parzen, F. (1962) On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33, 1065-76.
- روشی برای برآورد احتمال از نمونه‌ها.
- Pastore, R. E. and Farrington, S. M. (1996) Measuring the difference limen for identification of order of onset for complex auditory stimuli. *Perception & Psychophysics*, 58(4), 510-26.
- راجع به مبنای شنیداری کاربرد زبانی دمش به عنوان یک مشخصه تمایزدهنده.
- Pisoni, D. B. (1977) Identification and discrimination of the relative onset time of two-component tones: Implications for voicing perception in stops. *Journal of the Acoustical Society of America*, 61, 1352-61.
- بیشتر در مورد مبنای شنیداری کاربرد زبانی دمش به عنوان یک مشخصه تمایزدهنده.
- Rand, T. C. (1974) Dichotic release from masking for speech. *Journal of the Acoustical Society of America*, 55(3), 678-80.
- اولین نمونه اثر درک دوطرفه.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., and Carrell, T. D. (1981) Speech perception without traditional speech cues. *Science*, 212, 947-50.

اولین مقاله‌ای که نشان می‌دهد چگونه افراد جمله‌هایی را که منحصرأ با امواج سینوسی متغیربازمان سنتز شده‌اند، درک می‌کنند.

Rosenbium, L. D., Schmuckler, M. A., and Johnson, J. A. (1997) The McGurk effect in infants. *Perception & Psychophysics*, 59, 347-57.

Sekiyama, K. and Tohkura, Y. (1993) Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics*, 21, 427-44.

این نویسندگان دریافتند که اثر مک‌گرک برای افرادی که به زبان‌های متفاوتی صحبت می‌کنند، متفاوت است.

Shannon, C. E. and Weaver, W. (1949) *The Mathematical Theory of Communication*. Urbana: University of Illinois.

کتابی که «نظریه اطلاعات» را پایه‌ریزی کرد.

Shepard, R. N. (1972) Psychological representation of speech sounds. In E. E. David and p. B. Denes (eds.), *Human Communication: A unified view*. New York: McGraw-Hill, 67-113.

اندازه‌گیری فاصله در کی در یک ماتریس ابهام.

Walker, S., Bruce, V., and O'Malley, C. (1995) Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect. *Perception & Psychophysics*, 57, 1124-33.

نمونه‌ای جالب از اینکه چگونه دانش کل به جز می‌تواند سبب ایجاد اثر مک‌گرک شود.

Warren, R. M. (1970) Perceptual restoration of missing speech sounds. *Science*, 167, 392-3.

اولین نمونه از «اثر بازگردانی واج».

تمرین

واژگان تخصصی

این اصطلاحات را تعریف کنید: پیوستار محرکی، موج سینوسی گفتاری آنالوگ، درک دوطرفه، اثر مک‌گرک، اثر گانونگ، نسبت سیگنال به نوفه، ماتریس ابهام، مثلث‌بندی، مقیاس‌گذاری چندبعدی، وظیفه شناسایی، فاصله درکی.

پرسش

۱. کلمات "sue" و "see" را تولید و ضبط کنید. به طیف‌نگاشت آنها نگاه کنید. چقدر صداهای /s/ با یکدیگر تفاوت دارند؟ /s/ در کلمه "sue" را به /i/ در کلمه "see" و /s/ در کلمه "see" را به /u/ در کلمه "sue" بچسبانید. یک جبران هم‌تولیدی در اینجا عمل می‌کند. بستار هم‌تولیدی دخیل را شرح دهید.
۲. در یک سایت اینترنتی حاوی «اشعار اشتباه‌شنیده‌شده» مثل <http://www.kissthisguy.com> سه اشتباه درکی پیدا کنید که فکر می‌کنید می‌توانید آنها را بر اساس شباهت آوایی آکوستیکی بین پاره‌گفتاری که قصد تولید آن بوده، با اشتباه درکی فعلی توجیه کنید. نمره بیشتری می‌گیرید اگر طیف‌نگاشت اشعار شنیده‌شده و اشعار اصلی را که خودتان تولید کرده‌اید، بیاورید. آیا می‌توانید موردی از اشتباه درکی را پیدا کنید که شعری نباشد؟
۳. یک خط‌کش و یک پرگار برای رسم فضای درکی بر اساس ماتریس فاصله‌های زیر تهیه کنید. توجه کنید که ماتریس زیر، میزان تفاوت درک‌شده در صدای گویندگان را نشان می‌دهد. بدین‌صورت که ما جفت‌کلماتی را برای شنوندگان پخش کردیم و از آنها پرسیدیم «آیا دوباره صدای همان شخص است یا دو فرد متفاوتند؟» سپس شباهت را به‌صورت تعداد دفعاتی که گویندگان متفاوت، «یکسان» فرض شده‌اند، اندازه‌گیری کردیم. مقادیر فاصله سپس با استفاده از معادله (۵-۱) محاسبه شدند. شنوندگان پاسخ «یکسان» به

۱۵ درصد از مواردی دادند که صدای گوینده AJ با گوینده CN جفت شده بود $[1/9 = -\ln(0/15)]$. می‌توانید بگویید کدام دو گوینده دوقلو هستند؟

	AJ	CN	NJ	RJ
AJ		۱/۹	۰/۳	۱/۹
CN	۱/۹		۲/۳	۲/۵
NJ	۰/۳	۲/۳		۱/۹
RJ	۱/۹	۲/۵	۱/۹	

۴. فاصله درکی بین “d” و “θ” در جدول ۲-۵ را محاسبه کنید. آیا “θ” به “d” نزدیک‌تر است یا به “z”؟

۵. اگر توجه کرده باشید، من در این فصل دو علامت مختلف برای نشان‌دادن صداهای گفتاری به کار بردم. تعبیر این علامت‌ها که در متن تلویحاً به آنها اشاره شده بود، به‌صورت زیر است:

[θ] – تولید آوایی یا جنبه‌های فیزیکی و آکوستیکی صدا

“θ” – بازنمایی درکی صدا

برخی پژوهشگران استدلال می‌کنند که بازنمایی‌های درکی مثل “θ” برخاسته از بستارهای گفتاری هستند، بدین‌معنا که شنوندگان، گفتار را بر اساس فعالیت‌های دستگاه گفتار تعبیر می‌کنند تا منحصراً بر اساس الگوهای حسی. چه مباحثی در این فصل با این دیدگاه «بستارگرایی»^۱ درک گفتار مطابقت دارد یا ندارد؟

بخش ۲

تجزیه گفتار

فصل ۶

واکها

۶-۱. مدل‌های لوله‌ای تولید واکه

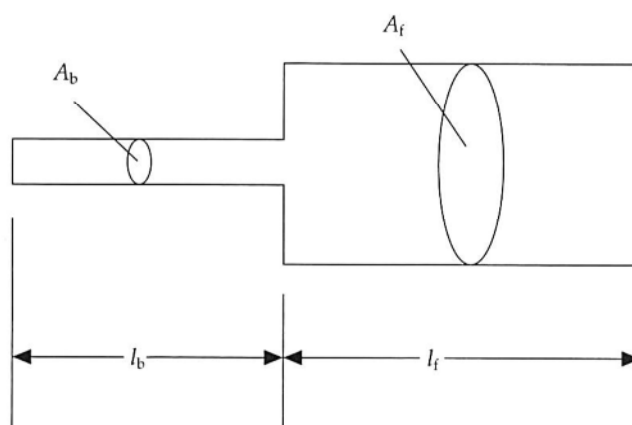
نظریه منبع-فیلتر پیش‌بینی فرکانس‌های بازخوانی دستگاه گفتار (مقادیر سازه‌ها) را به شرطی که سطح مقطع عرضی دستگاه گفتار مثل تولید شواً واحد باشد، میسر می‌سازد. علاوه بر آن، برای برخی وضعیت‌های تولیدی که سبب تغییر طول کلی دستگاه گفتار می‌شود (مثل پیش‌آمدگی لب‌ها و پایین‌رفتگی حنجره) نیز این تحلیل مناسب است. اما به هر حال، برای بررسی ویژگی‌های آکوستیکی سایر واکه‌ها لازم است این نظریه گسترش پیدا کند و شامل شکل‌بندی‌هایی^۱ از دستگاه گفتار شود که گرفتگی‌ها را نیز دربرگیرد.

یک راه برای مدل‌سازی تأثیرات آکوستیکی گرفتگی در تولید واکه این است که دستگاه گفتار را به صورت مجموعه‌ای از لوله‌ها و نه یک لوله در نظر بگیریم (فانت، ۱۹۶۰). در این بخش، تعدادی مدل لوله‌ای از دستگاه گفتار معرفی می‌شود که می‌تواند برای مدل‌سازی برخی از ویژگی‌های واکه‌ها به کار روند. بحث بر پایه

1. configuration

نظریات مطرح شده در فصل ۲ است، به ویژه بر پایه این واقعیت که فرکانس های بازخوانی یک لوله را می توان بر اساس طول آن محاسبه نمود.

در شکل ۶-۱، نمونه ای از یک مدل لوله ای که برای بعضی از شکل بندی های دستگاه گفتار مناسب است، نشان داده شده است (مبحث این بخش برگرفته از استیونس ۱۹۸۹ است). در این مدل، دستگاه گفتار به دو لوله تقسیم شده است. همان طور که شکل نشان می دهد، اگر سطح مقطع عرضی لوله جلویی A_f باشد و سطح مقطع عرضی لوله عقبی که خیلی کوچک تر از سطح مقطع عرضی لوله جلویی است، A_b باشد، می توان چنین فرض کرد که لوله عقبی در چاکنا بسته و در محل اتصال به لوله جلویی باز است. همچنین، می توان فرض نمود که لوله جلویی در محل اتصال به لوله عقبی بسته و در انتهای دیگر که لب ها هستند، باز است. بنابراین، از آنجا که هر دو لوله در یک انتها باز و در انتهای دیگر بسته هستند، می توان با استفاده از فرمول بازخوان های دستگاه گفتار ($F_n = (2n-1)c/4L$)، بازخوان های لوله ها (یا حفره های) جلو و عقب را جداگانه بر اساس طول آنها، l_b و l_f محاسبه نمود. بنابراین، بازخوان های حفره عقبی عبارت اند از: ($F_{bn} = (2n-1)c/4l_b$)، و بازخوان های حفره جلویی عبارت اند از: ($F_{fn} = (2n-1)c/4l_f$).



شکل ۶-۱. مدل دولوله ای که شبیه شکل دستگاه گفتار هنگام تولید [a] است.

شگردهای نظریه منبع-فیلتر

برای محاسبه طول دستگاه گفتار می‌توانید از نظریه منبع-فیلتر، و یا اگر خواستید، از سرعت صوت در گاز هلیوم استفاده کنید.

بر اساس نظریه منبع-فیلتر می‌دانیم که فرکانس‌های سازه‌ای واکه‌ها شوا با استفاده از طول دستگاه گفتار قابل محاسبه‌اند. بنابراین، اگر مقدار سازه‌ها را بدانید، می‌توانید طول دستگاه گفتار را تخمین بزنید. به‌عنوان مثال، اولین واکه کلمه *ahead* در یک دستگاه گفتار تقریباً فاقد گرفتگی (خنثی) تولید می‌شود، اما نوع دقیق این واکه دستگاه گفتار خنثی از شخصی به شخص دیگر متفاوت است. از این رو، بهتر است کمی تمرین کنید و ببینید واکه‌ای که شما با دستگاه گفتار خنثای خود تولید می‌کنید، چگونه است. برای این کار، لازم است طیف‌نگاشت تعداد متنوعی از واکه‌ها را، اغلب با کیفیتی بین [ə] و [æ] تهیه کنید و مقدار فرکانس سازه‌ها را یادداشت کنید (در این فصل به این بحث می‌پردازیم که چگونه شکل‌بندی‌های متفاوت دستگاه گفتار منجر به تولید مقادیر سازه‌ای متفاوت می‌شود). خواهید دید که وقتی سازه‌ها با فاصله مساوی از یکدیگر قرار گرفته‌اند، یعنی وقتی فاصله بین F_1 و F_2 به‌اندازه فاصله بین F_2 و F_3 است و به‌همین ترتیب الی آخر، دستگاه گفتار شما یک مقطع عرضی تقریباً واحد، یعنی شکل دستگاه گفتار خنثی را دارد.

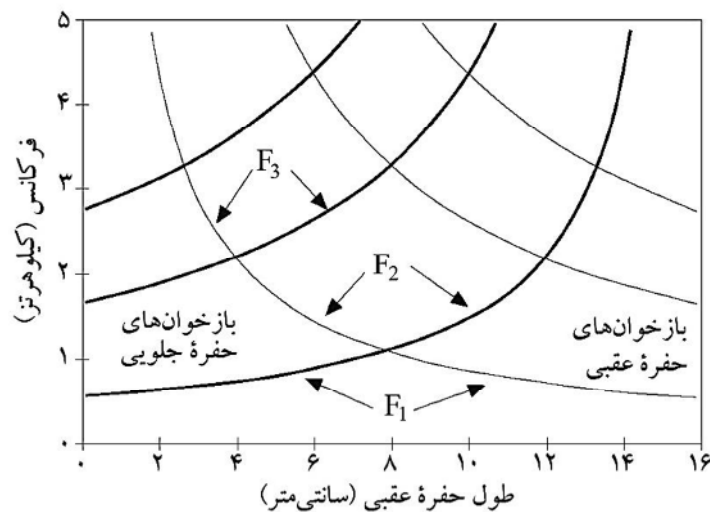
اکنون که واکه خنثای دستگاه گفتار خود را پیدا کرده‌اید، می‌توانید طول دستگاه گفتار خود را محاسبه کنید. با یک تغییر در فرمول سازه‌ها ($F_n = (2n-1)c/4L$) و قرار دادن L در سمت چپ ($L = (2n-1)c/4F_n$) مقدار طول را به دست آورید. به‌عنوان مثال، F_3 در واکه خنثای من ۲۶۰۰ هرتز است. بنابراین، طول دستگاه گفتار من $16/8$ سانتی‌متر است. روزی یک گزارشگر تلویزیون (که تاحدودی به لطافت صدای خود می‌بالید) از من پرسید: «این طیف‌نگاشت به شما چه چیزی را در مورد تفاوت صدای من با شما می‌گوید؟» وقتی به او گفتم طیف‌نگاشت نشان می‌دهد که دستگاه گفتار او کوچک‌تر از دستگاه گفتار من است، بسیار تعجب کرد. هرگز تصور نمی‌کردم که ممکن است نسبت به طول آن حساس باشد.

زمانی که یک فرد هلیوم استنشاق و سپس صحبت می‌کند، فرکانس‌های

بازخوانی دستگاه گفتار او افزایش پیدا می‌کند (و به‌همین دلیل صدای شخص خنده‌دار به نظر می‌رسد). لازم به گفتن نیست که دلیل آن، تغییر در طول دستگاه گفتار نیست، بلکه به‌دلیل تغییر رسانای آکوستیکی (هلیوم به جای هوا) و در نتیجه به‌دلیل تغییر در سرعت صوت است. از این رو، اگر طول دستگاه گفتار یک شخص را بدانید، باید بتوانید سرعت صوت در هلیوم را با اندازه‌گیری سازه‌های واکه شوا که با تنفس هلیوم تولید شده‌اند، به دست آورید. به‌عنوان مثال، اگر من در حالی که یک گاز ناشناخته را استنشاق می‌کنم، صحبت کنم، F_3 شوای من ۱۹۶۰ هرتز است. بنابراین، با استفاده از فرمول سازه‌ها ($c = F_n 4L / (2n-1)$) می‌توان سرعت صوت در گاز ناشناخته را حساب کرد که ۲۶۳۰۰ سانتی‌متر بر ثانیه است. اگر می‌خواهید بدانید این گاز چیست، آن را پیدا کنید.

در شکل ۲-۶ فرکانس‌های بازخوانی تولیدشده به‌وسیله این مدل بر اساس طول‌های مختلف حفره‌های جلو و عقب (با این فرض که طول کلی دستگاه گفتار ۱۶ سانتی‌متر است) نشان داده شده است. در این نمودار، پارامتر تولیدی بر روی محور افقی و برون‌داد آکوستیکی دستگاه گفتار بر روی محور عمودی نشان داده شده است. پارامتری که بر روی محور افقی شکل ۲-۶ نشان داده شده، طول حفره عقبی برحسب سانتی‌متر است و فرکانس‌های بازخوانی مدل لوله‌ای بر روی محور عمودی نشان داده شده‌اند. اگر حفره عقبی خیلی کوتاه باشد، فرکانس‌های بازخوانی آن خیلی بالا هستند که در نتیجه، پایین‌ترین بازخوان‌های دستگاه گفتار مربوط به حفره جلویی هستند (اگر $I_b=0$ باشد، مدل آن همانند مدل تک‌لوله‌ای واحدی است که در فصل دوم درباره آن بحث کردیم). وقتی طول حفره عقبی کمی بیش از ۴ سانتی‌متر باشد، پایین‌ترین بازخوان آن کمتر از دومین بازخوان حفره جلویی است. بنابراین، وقتی طول حفره عقبی بین ۴ تا ۸ سانتی‌متر باشد، پایین‌ترین بازخوان مدل لوله‌ای (F_1)، بازخوانی از حفره جلویی است، در حالی که دومین بازخوان (F_2).

یک بازخوان از حفره عقبی است. وقتی طول حفره عقبی برابر با طول حفره جلویی باشد ($l_b = l_f = 8 \text{ cm}$)، هر دو لوله فرکانس‌های بازخوانی یکسانی دارند. در یک سیستم دولوله‌ای مثل این، بازخوان‌های عقبی و جلویی با فرکانس یکسانی تولید نمی‌شوند، زیرا فرض اولیه ما که حفره عقبی در یک انتها باز و حفره جلویی در یک انتها بسته است، در واقع درست نیست. دو لوله، بسته به مقدار نسبی سطح مقطع عرضی خود (A_f و A_b) به لحاظ آکوستیکی با هم جفت می‌شوند و وقتی فرکانس‌های بازخوانی دو لوله به لحاظ نظری تقریباً مساوی باشد، اندکی با مقادیر پیش‌بینی شده برای لوله‌های جفت‌نشده تفاوت پیدا می‌کند (جفت‌شدگی آکوستیکی در فصل ۷ بیشتر مورد بحث قرار می‌گیرد). وقتی $l_b = l_f$ باشد، با جفت‌شدن آکوستیکی، مقدار F_1 مساوی ۷۵۰ هرتز و مقدار F_2 مساوی ۱۲۵۰ هرتز خواهد بود. اثر کامل جفت‌شدگی آکوستیکی به سطح مقطع عرضی دو لوله بستگی دارد و وقتی فرکانس بازخوان‌های دو حفره عقب و جلو به هم نزدیک می‌شوند، روی می‌دهد.



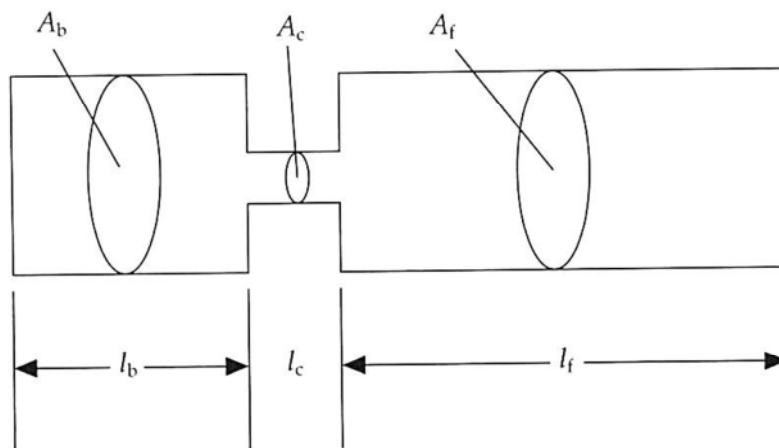
شکل ۶-۲. فرکانس‌های بازخوانی طبیعی لوله عقبی (خطوط کم‌رنگ) و لوله جلویی (خطوط پررنگ) برای مدل لوله‌ای شکل ۶-۱ بر حسب طول‌های مختلف حفره عقبی. طول کلی دستگاه گفتار ۱۶ سانتی‌متر است، بنابراین طول حفره جلویی ۱۶ سانتی‌متر منهای طول حفره عقبی است.

مدل لوله‌ای شکل ۱-۶ (که در آن طول حفره عقبی تقریباً به اندازه طول حفره جلویی است و سطح مقطع عرضی حفره عقبی خیلی کوچک‌تر از سطح مقطع عرضی حفره جلویی است) مدلی قابل قبول (هرچند بسیار ساده) از شکل دستگاه گفتار هنگام تولید [a] است. با استفاده از مدل دولوله‌ای شکل ۱-۶، به‌درستی پیش‌بینی می‌کنیم که واژه [a] فرکانس F_1 بالا و فرکانس F_2 پایین در مقایسه با F_1 و F_2 شوا دارد.

شکل ۳-۶ نوع دیگری از مدل لوله‌ای دستگاه گفتار را نشان می‌دهد. در این مدل، یک گرفتگی، لوله جلویی را از لوله عقبی جدا می‌کند. مثل مدل قبلی، می‌توان فرکانس‌های بازخوانی لوله‌های عقب و جلو را جداگانه با استفاده از همان فرمول‌هایی که برای محاسبه بازخوان‌های لوله واحد به کار رفت، به دست آورد. بازخوان‌های لوله جلویی با استفاده از فرمول بازخوان‌های لوله‌ای که در یک انتها بسته (گرفتگی) و در انتهای دیگر (لب‌ها) باز هستند، محاسبه می‌شوند. بازخوان‌های لوله عقبی با استفاده از فرمول بازخوان‌های لوله‌ای که در هر دو انتها بسته هستند (معادله ۱-۶) محاسبه می‌شوند، زیرا لوله عقبی در هر دو انتها عملاً بسته است. همان‌طور که در فصل ۲ گفتیم در این فرمول، مثل فرمول لوله‌ای که در یک انتها باز است، فرکانس‌های بازخوانی بر اساس طول لوله به دست می‌آیند و با کاهش طول لوله افزایش پیدا می‌کنند.

$$F_n = \frac{nc}{2L} \quad (1-6)$$

در این فرمول n شماره بازخوان، c سرعت صوت (۳۵۰۰۰ سانتی‌متر بر ثانیه) و L طول لوله (l_b) است.



شکل ۶-۳. مدل لوله‌ای شکل‌بندی‌هایی از دستگاه گفتار که یک گرفتگی کوچک در جایی در دستگاه گفتار دارند.

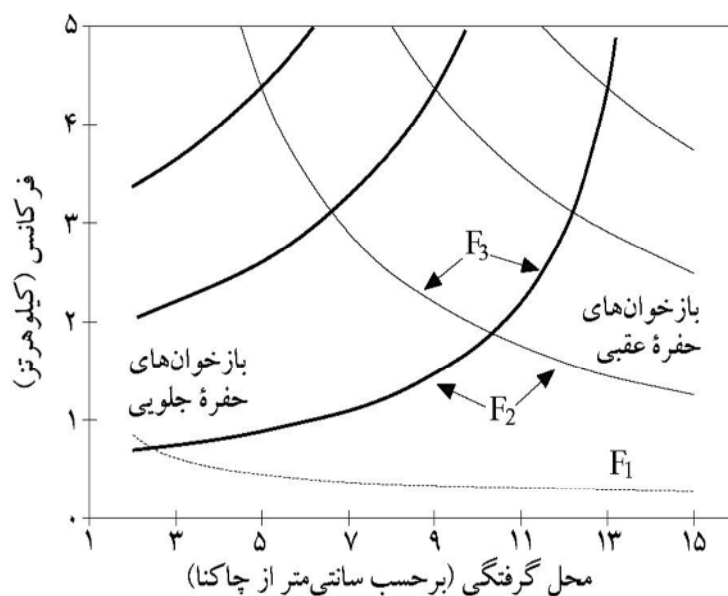
در این شکل‌بندی از دستگاه گفتار که در شکل ۶-۳ نشان داده شده، علاوه بر بازخوان‌های حفره‌های جلو و عقب، یک سیستم بازخوانی به نام بازخوان هلمهولتز^۱ به وسیله لوله عقبی و گرفتگی شکل می‌گیرد که در آن، حجم هوای محل گرفتگی مثل یک پیستون به درون و خارج گرفتگی نوسان می‌کند. وقتی بالای دهانه یک بطری خالی آبجو (یا نوشابه) می‌دمید، نوفه‌ای که تولید می‌شود، حاصل بازخوان هلمهولتز است (منبع صوتی آن از نوع اغتشاشی است). گردن بطری به مثابه گرفتگی و بدنه بطری به مثابه حفره عقبی دستگاه گفتار است. فرکانس بازخوانی طبیعی بازخوان هلمهولتز بر اساس حجم نسبی هوا در حفره عقبی و در گرفتگی تعیین می‌شود و با استفاده از معادله (۶-۲) می‌توان آن را محاسبه نمود.

$$f = \frac{c}{2\pi} \sqrt{\frac{A_c}{A_b l_b l_c}} \quad (۶-۲)$$

برون‌داد آکوستیکی این مدل لوله‌ای در شکل ۶-۴ نشان داده شده است. برای

1. Helmholtz resonator

محاسبه مقدار سازه‌ها در این نمودگرام، طول کلی دستگاه گفتار مانند مثال قبلی، ۱۶ سانتی‌متر در نظر گرفته شد. طول گرفتگی ۲ سانتی‌متر و طول لوله جلویی به طول لوله عقبی بستگی دارد ($l_f = 16 - 2 - l_b$). نمودگرام حاصل (شکل ۴-۶)، فرکانس‌های بازخوانی دستگاه گفتار را در محدوده‌ای از محل‌های گرفتگی که از چاکنا تا لب‌ها امتداد دارند (و بعضی از آنها به لحاظ تولیدی غیرممکن هستند)، را نشان می‌دهد. همان‌طور که در مورد شکل ۲-۶ گفته شد، وقتی لوله‌های جلویی و عقبی به لحاظ آکوستیکی جفت می‌شوند، فرکانس‌های بازخوانی در واقع یکدیگر را قطع نمی‌کنند (به عنوان مثال، مقدار F_2 وقتی $l_b = 10\text{cm}$ است). سازه اول (خط چین) که بازخوان هلمهولتز است، به عنوان تابعی از طول حفره عقبی تغییر می‌کند.



شکل ۴-۶. فرکانس‌های بازخوانی لوله عقبی (خطوط کم‌رنگ)، لوله جلویی (خطوط پررنگ) و بازخوان هلمهولتز (خط چین) برای مدل لوله‌ای که در شکل ۳-۶ نشان داده شد. ترسیم مقادیر فرکانس به عنوان تابعی از طول‌های متفاوت لوله عقبی (l_b) صورت گرفته است. طول گرفتگی ۲ سانتی‌متر و طول کلی مدل ۱۶ سانتی‌متر است.

برای فهمیدن آنچه در شکل ۴-۶ می‌بینید، به یاد آورید که فرکانس‌های بازخوانی یک لوله واحد به طول لوله بستگی دارند. لوله‌های کوتاه فرکانس‌های بازخوانی بالاتری نسبت به لوله‌های بلند دارند. بنابراین، با نگاه به مقادیر فرکانس F_2 ، می‌بینیم که وقتی گرفتگی در فاصله ۳ تا ۱۰ سانتی‌متری از چاکنا قرار دارد، دومین فرکانس بازخوانی (F_2) بالا در برون‌داد مدل، بازخوانی متعلق به لوله جلویی است، زیرا وقتی لوله جلویی بلند است، فرکانس‌های بازخوانی آن پایین است و وقتی لوله عقبی کوتاه است، فرکانس‌های بازخوانی آن بالاست. همان‌طور که گرفتگی به سمت جلو حرکت می‌کند، طول لوله عقبی بلندتر می‌شود. در نتیجه، فرکانس‌های بازخوانی آن پایین‌تر می‌روند و لوله جلویی کوتاه‌تر می‌شود. بنابراین، فرکانس‌های بازخوانی آن بالاتر می‌روند. وقتی گرفتگی کمی کمتر از ۱۱ سانتی‌متر از چاکنا فاصله دارد، حفره F_2 از جلو به عقب تغییر می‌کند. به عبارت دیگر، پایین‌ترین بازخوان حفره عقبی کمتر از پایین‌ترین بازخوان حفره جلویی می‌شود. در این مدل لوله‌ای، گرفتگی‌ای که ۱۱ سانتی‌متر از چاکنا فاصله داشته باشد، مانند شکل‌بندی دستگاه گفتار برای تولید واکه پیشین افراشته [i] است. چنان که در ادامه خواهیم دید، مقادیر سازه‌ای که مدل لوله‌ای پیش‌بینی می‌کند ($F_1=300$ هرتز، $F_2=1900$ هرتز، و $F_3=2200$ هرتز) تقریباً همان مقادیری هستند که در واکه [i] دیده می‌شود.

۶-۲. نظریه اختلال^۱

روش دیگری نیز برای مدل‌سازی پیامدهای آکوستیکی ایجاد گرفتگی در دستگاه گفتار وجود دارد که معمولاً به آن **نظریه اختلال** می‌گویند. در این روش، رابطه بین فشار هوا با سرعت نقش مهمی در مدل‌سازی ویژگی‌های آکوستیکی واکه‌ها ایفا می‌کند. (برای مطالب بیشتر در مورد نظریه اختلال رجوع کنید به چیبا و کاجی‌یاما^۲، ۱۹۴۱، و مرایاتی و همکاران^۳، ۱۹۸۸)

1. perturbation theory

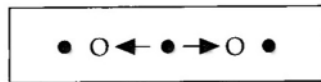
2. Chiba and Kajiyama

3. Mrayati et al

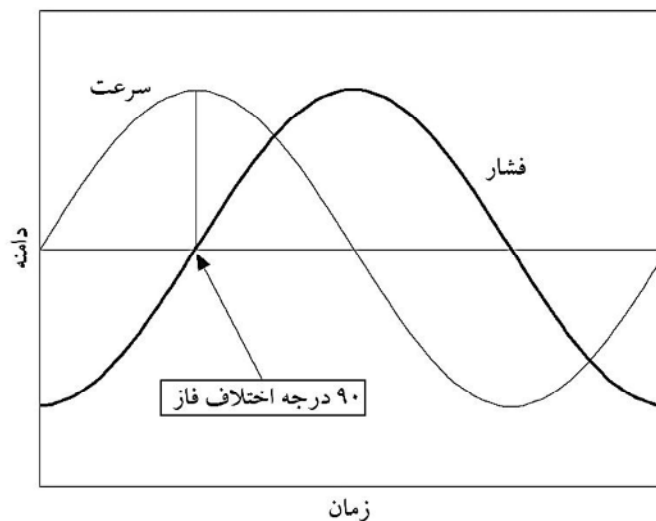
فرض کنید یک مولکول هوا مانند شکل ۵-۶ در جای خود در نوسان است. در لحظه‌ای که این ملکول به یکی از ملکول‌های مجاور خود نزدیک می‌شود، در مسیر تغییر جهت قرار می‌گیرد. در این لحظه، حرکت آن به دلیل تراکم با ملکول مجاور افت پیدا کرده و فشار هوا در حداکثر است، زیرا ملکول‌های هوا به هم فشرده شده‌اند. اما سرعت مولکول در حداقل است، زیرا در مرحله تغییر جهت قرار دارد. وقتی این مولکول به سوی دیگر تغییر جهت می‌دهد، در بین راه به نقطه‌ای مابین ملکول‌های مجاور خود می‌رسد که در این نقطه، تراکم (یا فشار هوا) در حداقل است، در حالی که سرعت در حداکثر است. این رابطه بین سرعت و فشار در شکل ۶-۶ نشان داده شده است. در نقاط حداکثر سرعت، فشار مساوی صفر است و در نقاط حداکثر فشار، سرعت مساوی صفر است.

شکل ۶-۷ که از چپا و کاجی‌یاما (۱۹۴۱) اقتباس شده، امواج ایستاده را در دستگاه گفتار (هم در برش عرضی آن و هم به صورت یک لوله مستقیم واحد) نشان می‌دهد. این شکل موج‌ها، جابه‌جایی ذره‌ای امواج ایستاده را (مثل شکل‌های ۲-۹ تا ۲-۷) نشان می‌دهند. اگر از فصل ۲ به یاد داشته باشید، گفتیم که بازخوان‌های طبیعی دارای حداکثر فشار در چاکنا و حداقل فشار در لب‌ها هستند. رابطه بین فشار و سرعت (شکل ۶-۶) سبب می‌شود که بازخوان‌های طبیعی دستگاه گفتار، همان‌طور که شکل ۶-۷ نشان می‌دهد، دارای حداقل سرعت در چاکنا و حداکثر سرعت در لب‌ها باشند. در شکل ۶-۷، نقاط حداکثر سرعت با v_n نشان داده شده است. این نقاط محل پادگره‌ها در موج ایستاده هستند. گره‌ها (نقاط حداکثر فشار) در سرعت صفر رخ می‌دهند. در این شکل، هم در تصاویر دستگاه گفتار و هم در لوله‌ها، امواج سرعت^۱ چهار فرکانس بازخوانی لوله واحدی که در یک انتها بسته است، با عنوان f_n نشان داده شده است. در هر یک از شکل‌ها، لوله واحد در بالا است و در پایین، مکان تقریبی گره‌های سرعت در دستگاه گفتار.

1. velocity wave



شکل ۶-۵. مولکول هوا در حال نوسان در جای خود. نقطه حداکثر فشار وقتی رخ می‌دهد که ملکول هوا به یکی از ملکول‌های مجاور خود نزدیک می‌شود، اما نقطه حداکثر سرعت وقتی روی می‌دهد که این ذره در فاصله یکسانی از ملکول‌های مجاور خود قرار می‌گیرد.

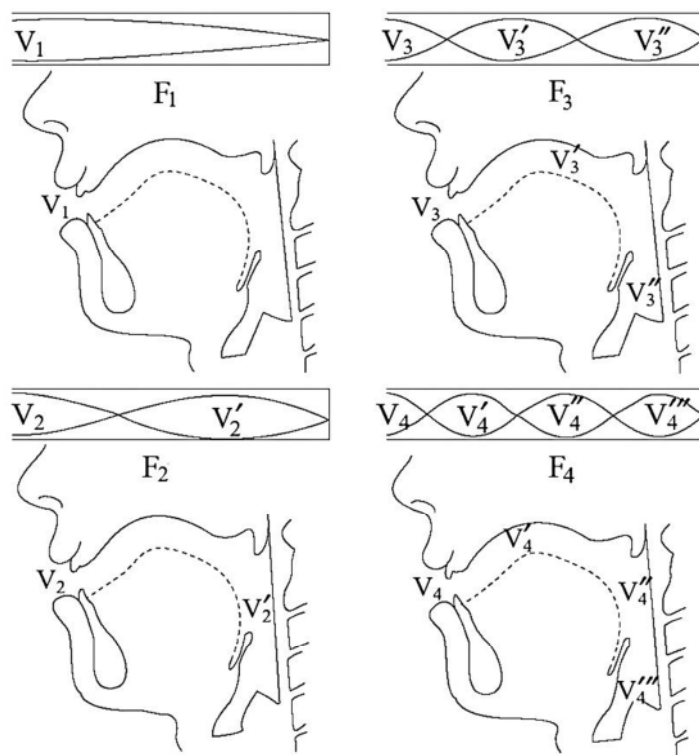


شکل ۶-۶. رابطه بین سرعت و فشار. وقتی سرعت در حداکثر مثبت یا منفی است، فشار مساوی صفر است و وقتی فشار در حداکثر مثبت یا منفی خود است، سرعت برابر صفر است.

نظریه اختلالی که چپا و کاجی‌یاما توصیف کرده‌اند، با در نظر گرفتن انرژی جنبشی موجود در نقاط دارای حداکثر سرعت و انرژی پتانسیل موجود در نقاط دارای حداکثر فشار، گرفتگی‌های دستگاه گفتار را به فرکانس‌های سازه‌ای مرتبط می‌کند. اگر در نقطه‌ای که انرژی جنبشی آن بالا (سرعت در حداکثر) است، در دستگاه گفتار گرفتگی ایجاد شود، حرکت ذره هوا با مقاومت روبرو می‌شود، در نتیجه فرکانس حرکت کاهش پیدا می‌کند؛ در حالی که از سوی دیگر، اگر در یک نقطه دارای انرژی پتانسیل بالا (فشار در حداکثر)، در دستگاه گفتار گرفتگی ایجاد شود، حرکت ذره هوا تقویت می‌شود، و در نتیجه فرکانس حرکت افزایش پیدا می‌کند. این

تأثیر گرفتگی‌ها بر فرکانس‌های بازخوانی دستگاه گفتار را می‌توان به‌صورت دو قاعده تجربی زیر خلاصه نمود:

۱. ایجاد گرفتگی در نزدیکی نقطه حداکثر سرعت (یا پادگره‌ها با علامت V_n)، فرکانس سازه‌ای را کاهش می‌دهد.
۲. ایجاد گرفتگی در نزدیکی نقطه حداکثر فشار (یا گره‌ها، جایی که خطوط یکدیگر را قطع می‌کنند)، فرکانس سازه‌ای را افزایش می‌دهد.



شکل ۶-۷. محل پادگره‌ها (نقاط حداکثر سرعت) و گره‌ها (نقاط حداکثر فشار) در موج ایستاده در لوله مستقیمی که در یک انتها باز است و در دستگاه گفتاری که فاقد گرفتگی است. پادگره‌ها با V_n (و از جلوی دستگاه گفتار با علامت پریم شماره‌گذاری شده‌اند) و گره‌ها با تقاطع امواج سینوسی در لوله مشخص شده‌اند. برای هریک از چهار سازه اول (با علامت F_n)، گره‌ها و پادگره‌ها در نمودارهای جداگانه نشان داده شده‌اند. برگرفته از چیپا و کاجی‌یاما (۱۹۴۱).

قواعد تجربی نظریه اختلال برای هریک از سازه‌ها (یا هریک از بازخوان‌های شکل ۶-۷) به‌طور جداگانه به کار می‌رود. به‌عنوان مثال، یک گرفتگی در حلق، در نزدیکی یک نقطه حداکثر فشار در بازخوان F_1 روی می‌دهد و بنابراین، طبق پیش‌بینی مدل اختلال، فرکانس F_1 در واکه [a] بالاتر از F_1 دستگاه گفتار در حالت خنثی (یا بدون گرفتگی) خواهد بود. در عین حال، یک گرفتگی در حلق در نزدیکی یک نقطه حداکثر سرعت در بازخوان F_2 نیز رخ می‌دهد و بنابراین، طبق پیش‌بینی مدل، F_2 واکه [a] پایین‌تر از F_2 دستگاه گفتاری خواهد بود که بدون گرفتگی است. در نتیجه، نظریه اختلال و مدل لوله‌ای که در شکل ۶-۱ نشان داده شد، نتیجه یکسانی برای [a] به دست می‌دهند.

سطح مقطع عرضی و گرفتگی دستگاه گفتار

در نظریه اختلال، پیش‌بینی فرکانس‌های سازه‌ای واکه با در نظر گرفتن تغییراتی در لوله واحد صورت می‌پذیرد. طبیعی است تغییر یک لوله واحد در دستگاه گفتار با ایجاد گرفتگی روی می‌دهد. از این رو، قواعد تجربی نظریه اختلال بر اساس محل گرفتگی بیان می‌شوند. دانستن این نکته مهم است که منظور از «گرفتگی» در اینجا، ضرورتاً فعال شدن یک بستر تنگ در دستگاه گفتار نیست، بلکه منظور از گرفتگی در اینجا، ایجاد یک سطح مقطع عرضی باریک در مقایسه با لوله واحد است. حتی اگر یک بستر گرفتگی فعال در دستگاه گفتار وجود نداشته باشد، ممکن است مساحت یک مقطع عرضی در یک نقطه خاص در دستگاه گفتار کوچک باشد. تولید واکه [æ] در کلمه *ash* نمونه‌ای از آن است. عکس‌برداری با اشعه X از برش عرضی دستگاه گفتار هنگام تولید [æ]، اثری از هیچ نوع گرفتگی فعال در دستگاه گفتار نشان نمی‌دهند. با این حال، با فرض سطح مقطع عرضی واحد در دستگاه گفتار، فرکانس F_1 و F_2 هر دو بالاتر از آن مقداری است که پیش‌بینی می‌شود. این الگوی آکوستیکی با دستگاه گفتاری مطابقت دارد که یک سطح مقطع عرضی کوچک دقیقاً بالای حنجره (که یک نقطه حداقل سرعت برای F_1 و F_2 است) دارد. «گرفتگی» در [æ] منحصراً یک ویژگی آناتومی دستگاه گفتار است، نه حاصل یک بستر گرفتگی فعال.

به وسیله دو مدل، نتیجه یکسانی برای F_2 واکه [i] نیز به دست می‌آید. نوموگرام شکل ۶-۴ نشان می‌دهد که مقدار F_2 برای یک گرفتگی که به فاصله ۱۰ سانتی‌متری از چاکنا قرار دارد، مثل واکه [i]، حدود ۲ کیلوهرتز است. این نتیجه مطابق با تحلیل نظریه اختلال از همین شکل‌بندی دستگاه گفتار است. با نگاهی دوباره به شکل ۶-۷، می‌بینیم که یک گرفتگی که در فاصله ۱۰ سانتی‌متری از چاکنا قرار دارد (فرض کنید طول کلی دستگاه گفتار حدود ۱۶ سانتی‌متر است)، دستگاه گفتار را در محلی که برای F_2 نقطه حداکثر فشار است، باریک می‌کند. بر اساس نظریه اختلال انتظار می‌رود که فرکانس F_2 بالاتر از آن مقداری باشد که در دستگاه گفتار بدون گرفتگی و خنثی است. به عنوان تمرین، پیش‌بینی و مقایسه سایر سازه‌های واکه [i] بر اساس دو نظریه اختلال و مدل لوله‌ای را بر عهده خواننده می‌گذارم.

پیش‌بینی سازه سوم (F_3) همخوان [ɪ] در انگلیسی آمریکایی، کاربردی‌بودن نظریه اختلال را به خوبی نشان می‌دهد. این آوا که با سه گرفتگی هم‌زمان در دستگاه گفتار (لبی، تیغه‌ای و حلقی) تولید می‌شود، با سایر آواهای واکه‌ای^۱ تفاوت دارد، زیرا فرکانس F_3 آن به طور غیرمعمول پایین است (لیندو^۲، ۱۹۸۵). شکل‌بندی دستگاه گفتار هنگام تولید [ɪ] در انگلیسی آمریکایی بسیار پیچیده‌تر از آن چیزی است که تاکنون در یک مدل لوله‌ای دیده‌ایم، اما پیش‌بینی F_3 پایین آن با استفاده از نظریه اختلال آسان است. جمع‌شدن^۳ یا برگشت^۴ زبان برای تولید [ɪ] (هیچ کدام فرقی ندارد)، دستگاه گفتار را در نقطه V'_3 تنگ می‌کند؛ گردی لب‌ها، در نقطه V_3 و گرفتگی حلقی در V''_3 . بنابراین، هر سه گرفتگی در [ɪ]، نزدیک نقاط حداکثر سرعت در موج ایستاده F_3 رخ می‌دهند و در نتیجه، نظریه اختلال به درستی پیش‌بینی می‌کند که فرکانس F_3 در [ɪ] در مقایسه با سایر آواهای واکه‌ای پایین خواهد بود.

مقایسه دو نظریه اختلال و مدل لوله‌ای برای واکه‌های [i] و [a] نشان می‌دهد که نظریه اختلال و مدل لوله‌ای پیش‌بینی‌های مشابهی برای فرکانس‌های سازه‌ای واکه‌ها

۱. vocalic. به واکه‌ها و همخوان‌های روان آواهای واکه‌ای می‌گویند. (مترجم)

2. Lindau

3. tongue bunching

4. tongue retroflexion

دارند. به هر حال، وقتی بیش از یک گرفتگی در دستگاه گفتار وجود دارد، مثل تولید [i]، کاربرد نظریه اختلال برای چنین شکل‌بندی خاصی در دستگاه گفتار راحت‌تر است، اما وقتی شکل‌بندی‌های دستگاه گفتار پیچیدگی کمتری دارند، استفاده از مدل‌های لوله‌ای برای پیش‌بینی کمی مقادیر سازه‌ها ساده‌تر است.

تفاوت مهم دیگری نیز بین نظریه اختلال و مدل لوله‌ای وجود دارد. در مدل لوله‌ای فرض بر این است که بازخوان‌های دستگاه گفتار برخاسته از ویژگی‌های بازخوانی حفره‌های خاصی در دستگاه گفتار هستند. به این معنا که مقادیر سازه‌ای که مدل‌های لوله‌ای پیش‌بینی می‌کنند، وابسته به یک یا چند لوله است. از این رو، فرضیه‌های مدل لوله‌ای خیلی بیشتر با تولیدهایی مطابقت دارد که گرفتگی‌های نسبتاً باریکی دارند و بنابراین بین لوله‌ها جفت‌شدگی خیلی زیادی روی نمی‌دهد. از سوی دیگر، فرضیه‌های نظریه اختلال خیلی بیشتر با تولیدهایی هماهنگ است که در آنها دستگاه گفتار عمدتاً بدون گرفتگی است.

۳-۶. واکه‌های مرجح^۱، نظریه ذره‌ای و نظریه پراکندگی تطبیقی^۲

وقتی در مورد مقادیر ذره‌ای در نگاشت پهنای چاکنا به برون‌داد آکوستیکی بحث می‌کردیم، شکلی که ارائه کردم (شکل ۲-۴) نوعی نوموگرام بود. مقادیر یک پارامتر تولیدی، یعنی پهنای چاکنا، بر روی محور افقی رسم شد و برون‌داد آکوستیکی فرضی بر روی محور عمودی. بر طبق پیش‌بینی نظریه ذره‌ای، در زبان‌های دنیا در نگاشت تولید به آکوستیک پهنای چاکنا، تولیدهای مرجح چاکنایی به‌وسیله نواحی باثبات ذره‌ای تعیین می‌شوند. بر همین اساس، استیونس (۱۹۸۹) اظهار می‌کند که نوموگرام‌هایی مانند نوموگرام‌های شکل ۲-۶ و ۴-۶ که تولید واکه‌های دهانی را به برون‌داد آکوستیکی آنها مرتبط می‌سازند، می‌توانند برای پیش‌بینی بعضی از روندهای جهانی که در فهرست واکه‌ها دیده می‌شود، به کار روند. وی خاطرنشان می‌کند که در نقاطی از نوموگرام که وابستگی حفره‌ای یک سازه تغییر می‌کند (نقاطی

1. preferred vowel

2. adaptive dispersion theory

که بازخوان‌های حفره عقب و جلو یکدیگر را قطع می‌کنند)، ممکن است گستره‌ای از محل‌های تولید، برون‌داد آکوستیکی تقریباً یکسانی داشته باشند. به‌عنوان مثال، در شکل ۴-۶، وقتی طول حفره عقبی بین ۱۰ تا ۱۱ سانتی‌متر است، یک ناحیه پایدار در F_2 ملاحظه می‌شود؛ طول‌های حفره عقبی در محدوده‌ای تقریباً برابر با ۲ سانتی‌متر، مقدار F_2 تقریباً یکسانی دارند. این مشاهده پیش‌بینی می‌کند که واکه [i]، که گرفتگی آن (یک گرفتگی کامی) در حدود ۱۰ سانتی‌متری از چاکنا قرار دارد و مشخصه آکوستیکی آن یک F_2 بالا و فاصله اندک بین F_2 و F_3 است، در زبان‌های دنیا متداول است. پیش‌بینی مشابهی می‌توان از نمودگرام شکل ۶-۲ در زبان‌های دنیا به دست داد. این شکل، تلاقی پایین‌ترین بازخوان‌های حفره عقب و جلو را وقتی طول لوله‌های عقب و جلو یکسان است (۸ سانتی‌متر در مدل)، نشان می‌دهد. این نوع گرفتگی دستگاه گفتار که معمولاً در تولید واکه [a] مشاهده می‌شود، مشخصه آکوستیکی آن یک F_1 با فرکانس نسبتاً بالا و یک F_2 با فرکانس نسبتاً پایین است. و در آخر اینکه، استیونس در نمودگرام‌هایی که در اینجا نیامده است، نشان می‌دهد که با افزودن گردی لب‌ها به مدل لوله‌ای شکل ۶-۳، یک ناحیه پایدار نزدیک نرم‌کام دیده می‌شود، همان‌طور که در تولید [u] روی می‌دهد.

بنابراین، نظریه ذره‌ای ادعا می‌کند که نواحی پایدار در نگاشت بین تولید و فرکانس F_2 ، پایدارترین واکه‌ها به‌لحاظ آکوستیکی، واکه‌های [i]، [a] و [u] هستند که در گوشه‌های دوزنقه واکه‌ها قرار دارند (یعنی واکه‌هایی که دارای بیشترین فضا برای تغییر محل تولید هستند). اینها واکه‌هایی هستند که به کرات در زبان‌های دنیا رخ می‌دهند (مدیسون^۱، ۱۹۸۴).

نظریه پراکندگی تطبیقی لیندبلوم^۲ (۱۹۹۰) تبیین متفاوتی برای مرجح‌بودن واکه‌های گوشه‌ای در زبان‌ها دارد. در رویکرد پراکندگی تطبیقی، واکه‌های گوشه‌ای متداول‌ترین واکه‌ها در زبان‌های دنیا هستند؛ دقیقاً به این دلیل که این واکه‌ها در گوشه‌های دوزنقه واکه‌ها قرار دارند. به‌این‌معنی که اگر محدوده مقادیر ممکن F_1 و

1. Maddieson

2. Lindblom

F_2 را که در واکه‌ها می‌تواند تولید شود، در نظر بگیریم، آن واکه‌هایی که دارای حداکثر تمایز هستند، با اعتبار بیشتری از یکدیگر تشخیص داده می‌شوند. بنابراین، اگر چنین فرض کنیم که توانایی شنوندگان در شنیدن تمایزهای واکه‌ای، فشار گزینشی بر فهرست واحدهای آوایی وارد می‌کند (در تحول در زمانی زبان)، می‌توانیم پیش‌بینی کنیم که متداول‌ترین واکه‌ها در زبان‌های دنیا آنهایی هستند که بالاترین درجه مقادیر سازه‌ای را دارند (بخش ۶-۵ را ببینید). نظریه پراکندگی تطبیقی با در نظر گرفتن نقش شنونده، نظریه‌ای درباره پایداری در ارتباط گفتاری است؛ در حالی که نظریه ذره‌ای، نظریه‌ای درباره پایداری، اما تنها در یک جنبه ارتباط یعنی نگاشت تولید-به-آکوستیک است. اگرچه نظریه پراکندگی تطبیقی پیش‌بینی‌هایی در مورد واکه‌های مرجح دارد، اما این نظریه به سایر آواهای گفتاری تعمیم نیافته است. از سوی دیگر، در حالی که نظریه ذره‌ای بر روی یک جنبه کوچک‌تر ارتباط گفتاری تمرکز دارد، اما برای واحدهای گفتاری مختلفی به کار رفته است.

۶-۴. سازه‌ها و فضای واکه‌ای آکوستیکی^۱

شکل ۶-۸ طیف‌نگاشت کلماتی از زبان مزاتک جالاپا^۲ (یکی از زبان‌های خانواده زبانی اوتومانگوان^۳ که در ایالت واکاکا^۴ در کشور مکزیک به آن صحبت می‌شود) را نشان می‌دهد. این کلمات که با واکه‌های واک‌دار ساده^۵ در این زبان تولید شده‌اند، به صورت زیر هستند:

[si] کثیف

[se] او آواز می‌خواند

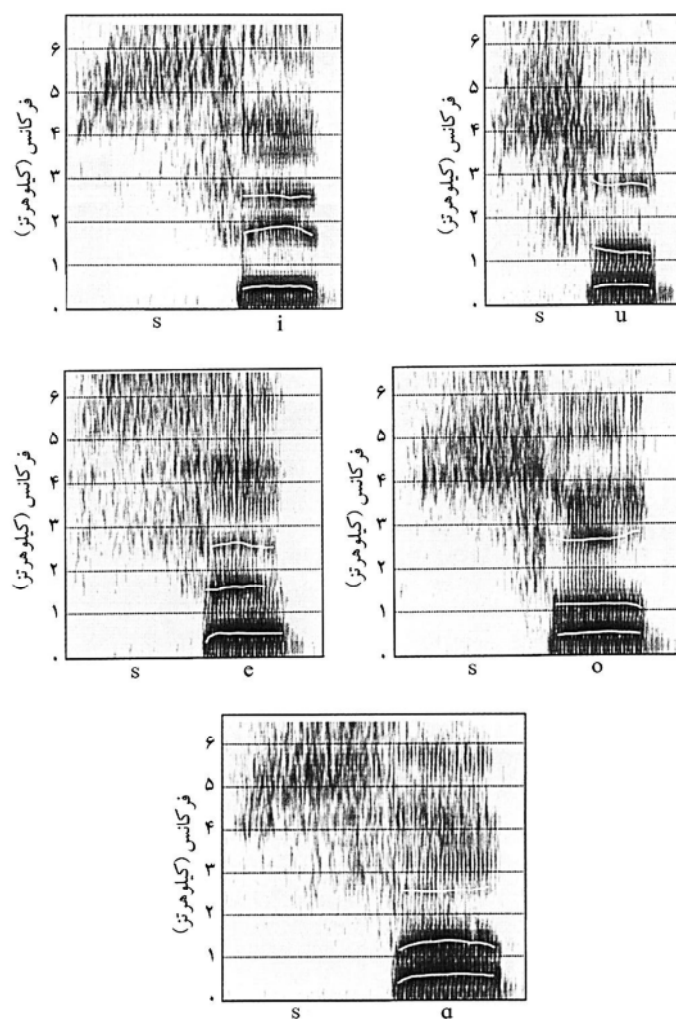
[sq] ماه

[so] شما

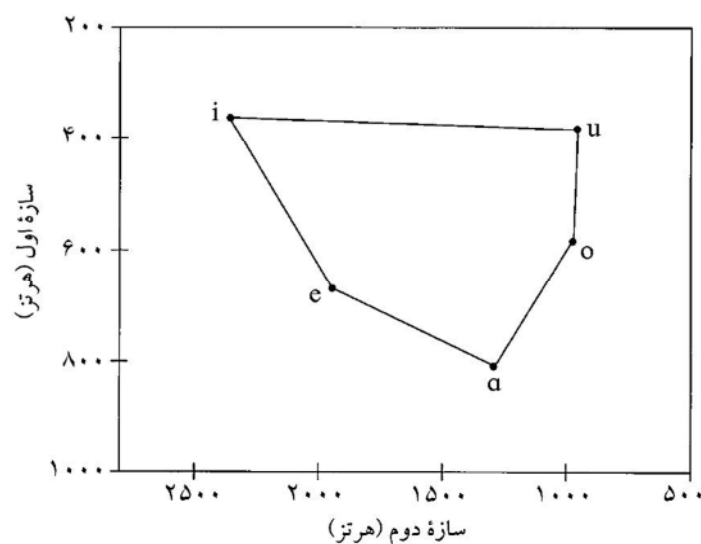
[su] ولرم

1. acoustic vowel space 2. Jalapa Mazatec 3. Otomanguen 4. Oaxaca
 ۵. plain voiced vowel. در بسیاری از زبان‌های بومی آمریکا، تقابل واک‌دار-بی‌واک در واکه‌ها نیز مشاهده می‌شود. (مترجم)

در طیف‌نگاشت‌ها، فرکانس مرکزی سه سازه اول برای هر واکه مشخص شده است. همان‌طور که می‌بینید، سازه‌ها در طول این واکه‌های ساده نسبتاً ایستا هستند. علاوه بر این، طیف‌نگاشت‌ها نشان می‌دهند که تمایزات واکه‌ای به وسیله فرکانس دو سازه اول و دوم منتقل می‌شود.



شکل ۶-۸. طیف‌نگاشت کلمات زبان مزاتک جالاپا که توسط یک گویشور مرد تولید شده‌اند. آوانویسی کلمات بر اساس آی‌پی‌ای در زیر هر طیف‌نگاشت آمده است.

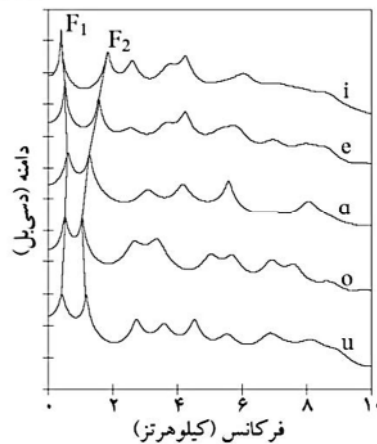


شکل ۹-۶. فضای آکوستیکی واکه‌های واکن‌دار ساده در زبان مزاتک جالاپا، میانگین چهار گویشور مرد.

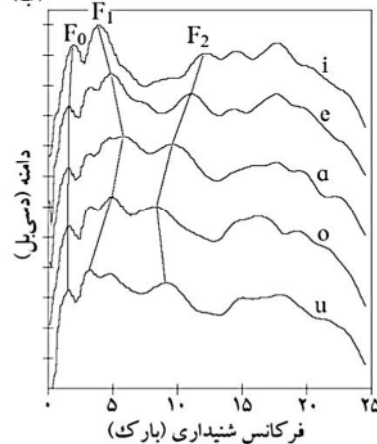
در شکل ۹-۶، نمودار میانگین فرکانس F_1 و F_2 در پنج واکه زبان مزاتک به صورت یک «فضای واکه‌ای آکوستیکی» نشان داده شده است. توجه کنید که فضای واکه‌ای آکوستیکی شبیه به مثلث سنتی شهودی واکه‌هاست. ارتفاع واکه با فرکانس F_1 همبستگی منفی دارد؛ F_1 واکه‌های [افراشته] پایین و F_1 واکه‌های [افتاده] بالاست (برای تأکید بر همبستگی بین ارتفاع شهودی واکه با فرکانس F_1 ، فرکانس‌های پایین در بالای محور F_1 قرار گرفته‌اند). به همین صورت، پیشین‌بودن واکه با F_2 همبستگی دارد؛ واکه‌های [پیشین] F_2 بالا، و واکه‌های [پسین] F_2 پایین دارند. این همبسته‌های آکوستیکی مشخصه‌های واکه برای قرن‌ها مورد توجه بوده است و به‌ویژه وقتی طیف‌نگار صوتی به وجود آمد، آشکار شدند (جوس، ۱۹۴۸). به نظر می‌رسد که مشخصه‌های تمایزدهنده واکه‌ها به این ویژگی‌های آکوستیکی مربوط است و نه به ویژگی‌های تولیدی، زیرا در تولید واکه تفاوت‌های فردی وجود دارد. علاوه بر آن، بین الگوهای زبان‌شناختی «افراشتگی» و «پیشین‌بودن» واکه و مقادیر اندازه‌گیری‌شده ارتفاع و پیش‌بودن زبان در طی تولید واکه (جانسون و همکاران، ۱۹۹۳؛ لدف‌وگد و

همکاران، ۱۹۷۲) و نیز بین میزان کشش ماهیچه‌ای اندازه‌گیری شده و «سختی»^۱ زبان‌شناختی (رافائل و بل‌پرتی^۲، ۱۹۷۵؛ لندو، ۱۹۷۸، ۱۹۷۹؛ هله و استیونس^۳، ۱۹۶۹؛ پرکل^۴، ۱۹۷۱؛ استوک‌ول^۵، ۱۹۷۳) تناقض ملاحظه می‌شود.

(الف)



(ب)



شکل ۶-۱۰. طیف آکوستیکی و شنیداری واکه‌هایی که در شکل ۶-۸ نشان داده شد: (الف) طیف‌های ال پی‌سی که از میان واکه گرفته شده‌اند، (ب) طیف‌های شنیداری همان پنجره‌های شکل موج آکوستیکی. بر روی محور عمودی، فاصله بین علامت‌گذاری‌ها ۲۰ دسی‌بل است. طول پنجره برای هر دو طیف ال‌پی‌سی و شنیداری ۲۰ میلی‌ثانیه است.

- | | |
|----------------------|---------------------------|
| 1. tenseness | 2. Raphael and Bell-Berti |
| 3. Halle and Stevens | 4. Perkell |
| | 5. Stockwell |

۵-۶. بازنمایی شنیداری و آکوستیکی واکه‌ها

روابط غیرخطی سیستم شنوایی که در فصل ۳ درباره آنها بحث کردیم، سبب می‌شود که مشخصه‌های شنیداری واکه‌ها از بازنمایی آکوستیکی آنها که در طیف‌نگاشت قابل مشاهده و به وسیله نظریه منبع-فیلتر قابل پیش‌بینی است، متفاوت باشد (برای مطالب بیشتر در زمینه بازنمایی شنیداری واکه‌ها رجوع شود به لی‌لین کرائنس و لیندبلوم^۱، ۱۹۷۲؛ سیردال و گوپهال^۲، ۱۹۸۶؛ میلر، ۱۹۸۹؛ تران‌مولر^۳، ۱۹۸۱).

شکل ۱۰-۶ بعضی از تفاوت‌های بین طیف‌های آکوستیکی با طیف‌های شنیداری را نشان می‌دهد: شکل (الف) طیف‌های ال‌پی‌سی پنج واکه زبان مزاتک را که در شکل ۸-۶ نشان داده شد، نشان می‌دهد و (ب) طیف‌های شنیداری همان واکه‌ها را. این طیف‌های شنیداری با استفاده از همان پنجره‌های شکل موج آکوستیکی که برای تولید طیف‌های ال‌پی‌سی به کار رفتند، با شبیه‌سازی رایانه‌ای سیستم شنوایی بیرونی (بخش ۵-۴) تولید شده‌اند. دو سازه اول واکه‌ها را در طیف‌های آکوستیکی با دو خط به هم وصل کردم و در طیف‌های شنیداری با یک خط دیگر، اولین همساز واکه‌ها را به هم وصل نمودم.

طیف‌های آکوستیکی و شنیداری به لحاظ محدوده فرکانسی (هر دو از صفر تا ۱۰ کیلوهرتز) با هم مطابقت دارند. از این رو، تفاوت‌هایی که در بازنمایی آنها می‌بینید، مربوط به ویژگی‌های تجزیه آکوستیکی و شنیداری است، نه به روش بازنمایی. یکی از تفاوت‌های چشمگیر بین طیف‌های شنیداری و آکوستیکی در نسبت فضای اشغال‌شده به وسیله F_1 و F_2 است. در طیف‌های آکوستیکی، F_1 و F_2 محدود به یک‌چهارم پایینی مقیاس فرکانسی هستند، در حالی که در طیف‌های شنیداری، نیمی از مقیاس فرکانسی را دربرمی‌گیرند. در نتیجه، تغییرات F_1 و F_2 به سادگی در طیف‌های شنیداری قابل مشاهده است، در حالی که در طیف‌های ال‌پی‌سی تا حدی مبهم باقی می‌ماند. این تفاوت مطابق با این مشاهده است که بیشتر اطلاعات مورد استفاده در ارتباط گفتاری روی فرکانس‌های کمتر از ۲۰۰۰ هرتز موجود است (میلر و نایسلی ۱۹۵۵).

1. Liljencrants and Lindblom

2. Syrdal and Gopthal

3. Traunmüller

یکی دیگر از تفاوت‌های بین طیف‌های شنیداری با آکوستیکی واکه در شکل ۶-۱۰، این است که اولین همساز (F_0) در طیف شنیداری تفکیک شده، اما در طیف ال‌پی‌سی چنین نشده است. البته، می‌توان با محاسبه طیف‌های اف‌اف‌تی، مقادیر F_0 را در بازنمایی آکوستیکی نیز تفکیک نمود، اما در این صورت، قله‌های سازه‌ای به‌خوبی مشخص نمی‌شوند. طیف‌های شنیداری جالب‌اند، زیرا پهنای باند باند‌های بحرانی شنیداری در فرکانس‌های پایین به‌اندازه کافی باریک است که همساز اول به‌عنوان یک قله مجزا تفکیک شود و در فرکانس‌های بالاتر، به‌اندازه کافی بزرگ است که همسازهای مجاور به‌صورت قله‌های سازه‌ای در یکدیگر محو شوند. بنابراین، پاسخ فرکانسی غیرخطی سیستم شنوایی اهمیت نسبی ناحیه F_1/F_2 را تقویت می‌کند، در حالی که پهنای باند باند‌های بحرانی، تمایل به محو همسازها در یکدیگر به‌صورت قله‌های سازه‌ای دارد.

۶-۶. درک زبان‌گذر واکه

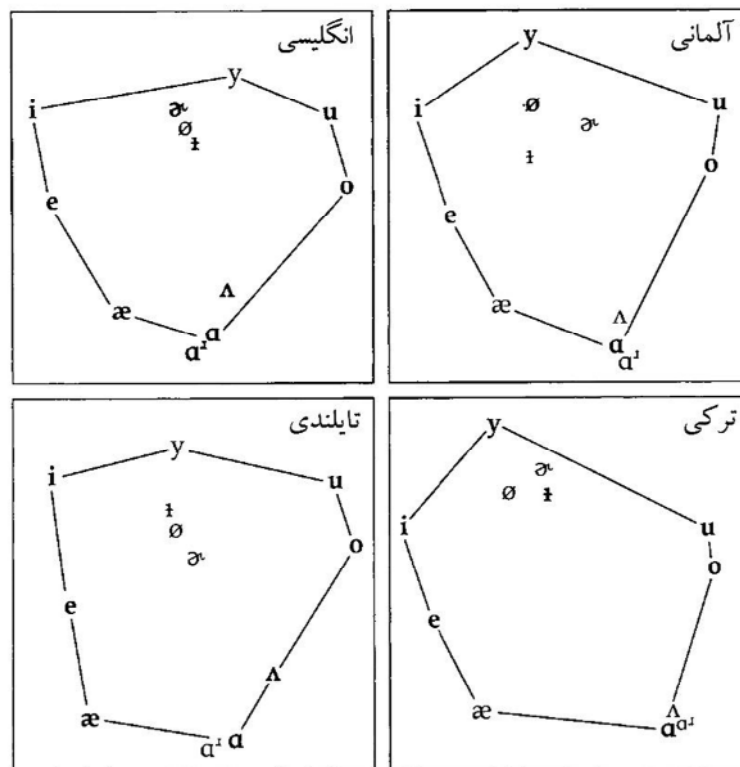
شکل ۶-۱۱ فضاهای درکی را نشان می‌دهد که بر مبنای مطالعه زبان‌گذر تربیک^۱ (۱۹۷۷) در مورد درک واکه‌ها رسم شده‌اند. تربیک واکه‌های ساده‌ای را که خودش در بافت [bɒb-] تولید کرده بود، به کار برد. او یک روش واقعاً جالب برای اندازه‌گیری فاصله درکی به کار برد. بدین صورت که واکه‌ها را در دسته‌های سه‌تایی برای شنوندگان پخش کرد و از آنها خواست قضاوت کنند که کدام دو واکه بیشترین شباهت را به یکدیگر دارند و کدام دو بیشترین تفاوت را. به‌عنوان مثال، شنونده ممکن است واکه‌های زیر را بشنود و [i] و [u] را متفاوت‌ترین و [y] و [u] را شبیه‌ترین معرفی کند:

A	B	C
bəbi	bəby	bəbu

او بدین ترتیب یک «ماتریس عدم شباهت»^۲ از تعداد دفعاتی که یک جفت معین متفاوت‌تر از جفتی دیگر قضاوت شده بود، به دست آورد.

1. Terbeek

2. dissimilarity matrix



شکل ۶-۱۱. فضاهای واکه‌ای در کی در شنوندگانی که به زبان‌های انگلیسی، آلمانی، تایندی و ترکی صحبت می‌کنند. این فضاها بر اساس ماتریس‌های عدم شباهت تریبک^۱ (۱۹۷۷) (پیوست I) محاسبه شده‌اند. من فضاها را با استفاده از ضریب رگرسیون که در روش مقیاس‌گذاری چندبعدی (ام‌دی‌اس) عدم شباهت اولیه را به فاصله‌ی نرمال‌شده ربط می‌دهد، رسم کردم. سپس فضاها را به‌نحوی چرخاندم که خطی که [i] را به [u] وصل می‌کند، افقی باشد. واکه‌هایی که کم‌رنگ‌تر نشان داده شده‌اند، در زبان مورد نظر تولید نمی‌شوند (به‌عنوان مثال، در انگلیسی واج /y/ وجود ندارد).

من ماتریس‌های عدم شباهت تریبک در مورد شنوندگان چهار زبان متفاوت را انتخاب کرده و نقشه‌های درکی را که در شکل ۶-۱۱ می‌بینید، رسم کردم. نکات جالب زیادی در این نقشه‌ها وجود دارد، اما در اینجا فقط به چهار مشاهده‌ی کلی در مورد درک واکه می‌پردازیم.

1. Terbeek

اول، توجه کنید که چقدر این فضاهاى واكه‌اى دركى به فضاى واكه‌اى آکوستيکى در شکل ۶-۹ شبیه هستند. این شباهت بسیار قابل توجه است، زیرا فضاهاى دركى بر اساس قضاوت شنوندگان در مورد شباهت و عدم شباهت رسم شده‌اند، در حالى که فضاى آکوستيکى فقط نمودار F_1 و F_2 واكه‌هاست. واضح است که فرکانس‌هاى سازه‌اى در درک واكه اهميت دارند.

دوم، توجه کنید که چگونه فهرست آواهاى تمایزدهنده در یک زبان بر درک تأثیر دارد. برای مثال، فضاهاى آلمانى و ترکی را با فضاهاى انگلیسى و تایلندى مقایسه کنید. در هر دو زبان آلمانى و ترکی، /y/ به‌عنوان یک واج به کار مى‌رود. بنابراین، جالب است که [y] از سایر واكه‌ها اندكى فاصله مى‌گیرد و به‌صورت یک قله در بالای فضاى دركى قرار مى‌گیرد. از سویی واكه [ʌ] نه در آلمانى و نه در ترکی به‌عنوان یک آواى تمایزدهنده وجود ندارد و به نظر مى‌رسد که در هر دوی این زبان‌ها در مقایسه با تایلندى و انگلیسى، واكه [ʌ] با واكه [a] ادغام شده است. اگر یک آوا در زبانی به‌صورت تمایزدهنده به کار رود، شنوندگان مایل‌اند آن را متفاوت‌تر از آواهاى مجاور بشنوند در مقایسه با وقتی که آن آوا برای تمایز واژگانی در زبان مذکور به کار نمى‌رود.

سوم، ببینید در مورد [æ] چگونه است. اگرچه این آوا در آلمانى و ترکی تمایزدهنده نیست، اما این شنوندگان [æ] را از [e] و [a] متمایز مى‌کنند، همان‌طور که شنوندگان انگلیسى و تایلندى که برای آنها [æ] یک واج است، این کار را مى‌کنند. بر اساس این مشاهده به نظر مى‌رسد که تمایز خام آکوستيکى یا شنیدارى در همه زبان‌ها اهميت دارد. واكه [ʌ] ممکن است با [a] ادغام شود، زیرا به‌لحاظ شنیدارى به آن شبیه است، اما [æ] آنقدر متمایز است که در برابر ادغام با سایر واكه‌ها مقاومت مى‌کند.

در نهایت، توجه کنید که چگونه هندسه فضاى دركى از زبانی به زبان دیگر تغییر مى‌کند. این آواها سیگنال‌هاى آکوستيکى کاملاً یکسانی هستند که توسط شنوندگانی با پیشینه زبانی متفاوت شنیده شده‌اند. بنابراین، جالب است که برای

گویشوران انگلیسی زبان، واکه گرد پیشین افراشته [y] به واکه گرد پسین افراشته [u] شبیه تر است، در حالی که برای گویشوران آلمانی و ترکی، [y] به واکه گسترده پیشین افراشته [i] شباهت بیشتری دارد. این تفاوت زبان گذر ممکن است به اینکه چگونه گویشوران انگلیسی آمریکایی [u] را تلفظ می کنند، مربوط باشد. به عبارت دیگر، گویشوران انگلیسی بیشتر تمایل دارند این آوا را پیشین تولید کنند که در نتیجه، بیشتر شبیه واکه گرد مرکزی افراشته [ɥ] یا حتی واکه گسترده و مرکزی [i] تولید می شود. هر دوی این تولیدها سبب می شود سازه دوم با فرکانس بالاتر همانند سازه دوم [y] تولید شود. بنابراین، هرچند که فرکانس سازه دوم واکه [u] در مطالعه تربیک ممکن است پایین بوده باشد، اما برای گویشوران انگلیسی، معرف یک مقوله واکه ای است که می تواند F_2 آن با فرکانس بالاتر نیز تولید شود و در نتیجه، [y] شباهت بیشتری به [u] دارد تا به [i]. برای گویشوران ترکی و آلمانی ممکن است درک [y]، یادآور جایگاه پیشین زبان برای تولید [i] باشد. بنابراین، شباهت نسبتاً زیاد [i] به [y] منعکس کننده دانش تولیدی است که گویشوران انگلیسی فاقد آن هستند. نوع مشابهی از نظام درکی زبان ویژه را می توان در پسین بودن نسبی [a] در ترکی در مقایسه با جایگاه مرکزی تر آن در سایر زبان ها نیز دید. ممکن است وجود یک قاعده هماهنگی واکه ای^۱ در مشخصه پسین در زبان ترکی، سبب افزایش شباهت زبان شناختی [a] با سایر واکه های [پسین] برای گویشوران این زبان شود؛ مسئله ای که در این مطالعه به واج شناسی آن در سایر زبان ها توجه نشده است.

1. vowel harmony

منابعی برای مطالعه بیشتر

- Chiba, T. and Kajiyama, M. (1941) *The Vowel: Its nature and structure*, Tokyo: Kaiseikan.
- کتابی کلاسیک که به مباحث گسترده‌ای در زمینه آواشناسی از جمله به آکوستیک و درک واکه‌ها می‌پردازد. این کتاب بیشتر به دلیل مبحث قوی و مؤثر مدل اختلال در آکوستیک واکه‌ها شناخته شده است.
- Hagiwara, R. (1995) Acoustic realizations of American /r/ as produced by women and men. *UCLA Working Papers in Phonetics*, 90, 1-187.
- مطالعه‌ای در مورد آکوستیک /r/ و آناتومی دستگاه گفتار.
- Halle, M. and Stevens, K. N. (1969) On the feature "advanced tongue root." *Quarterly Progress Report*, 94, 209-15. Research Laboratory of Electronics, MIT.
- Johnson, K., Ladefoged, P., and Lindau, M. (1993) Individual differences in vowel production. *Journal of the Acoustical Society of America*, 94, 701-14.
- Ladefoged, P., DeClerk, J., Lindau, M., and Papcun, G. (1972) An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics*, 22, 48-75.
- Liljencrants, J. and Lindblom, Björn (1972) Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48, 839-62.
- مطالعه‌ای مؤثر در زمینه پراکندگی واکه با این رویکرد که پراکندگی واکه ممکن است به الگوهای زبان‌گذر در سیستم‌های واکه‌ای مربوط باشد.
- Lindau, M. (1978) Vowel features. *Language*, 54, 541-63.
- یک مطالعه مهم در آواشناسی واکه.
- Lindau, M. (1979) The feature "expanded." *Journal of Phonetics*, 7, 163-76.
- Lindau, M. (1985) The story of /r/. In V. Fromkin (ed.), *Phonetic Linguistics: Essays in honor of Peter Ladefoged*, Orlando, FL: Academic Press.
- بررسی برخی ویژگی‌های تولیدی و آکوستیکی /r/ در زبان‌ها.
- Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H&H theory. In w. j. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modeling*, Dordrecht: Kluwer, 403-39.
- به معرفی نظریه‌ای در زمینه تنوع آوایی می‌پردازد. براساس این نظریه، گویشوران فعالانه گونه‌های آوایی را با توجه به نیازهای ارتباطی شنونده تغییر می‌دهند. این نظریه همچنین منجر به پیدایش نظریه «پراکندگی تطبیقی» در رده‌شناسی زبان‌گذر واکه‌ها شد.
- Maddieson, I. (1984) *Patterns of Sounds*, Cambridge: Cambridge University Press.
- یک مطالعه بسیار تأثیرگذار در رده‌شناسی آوایی به همراه فصولی راجع به انواع اصلی آواهای گفتاری.
- دادگان واحدهای آوایی که برای این مطالعه گردآوری شد، نقطه عطفی در پژوهش‌های آواشناختی محسوب می‌شود.
- Miller, J. D. (1989) Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America*, 85, 2114-34.
- مدلی از بازنمایی شنیداری واکه.
- Mrayati, M., Carré, R., and Guérin, B. (1988) Distinctive regions and modes: A new theory of speech production *Speech Communication*, 7, 257-86.
- Pant, G. (1960) *Acoustic Theory of Speech Production*, The Hague: Mouton.
- مدل‌های لوله‌ای بسیار متنوع فانت برای واکه‌ها (از جمله مدل‌های لوله‌ای ساده مستقیمی که در این فصل به کار رفت)، استاندارد مدل‌سازی آکوستیکی واکه‌ها را تعیین کرد.

- Perkell, J. (1971) Physiology of speech production: A preliminary study of two suggested revisions of the features specifying vowels. *Quarterly Progress Report*, 102, 123-39. Research Institute of Electronics, MIT.
- Raphael, L. J. and Bell-Berti, F. (1975) Tongue musculature and the feature of tension in English vowels. *Phonetica*, 32, 61-73. Are [tense] vowels "tense"?
- Stevens, K. N. (1989) On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.
- نشان می‌دهد چگونه می‌توان نوموگرام‌ها را برای شناسایی نواحی ذره‌ای واکه به کار برد.
- Stockwell, R. p. (1973) Problems in the interpretation of the Great English Vowel Shift. In M. E. Smith (ed.), *Studies in Linguistics in Honor of George L. Trager*, The Hague: Mouton, 344-62.
- تحلیل زبان‌شناختی بلند/کوتاه و سخت/نرم در زبان انگلیسی.
- Syrdal, A. K. and Gophal, H. S. (1986) A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79, 1086-1100.
- مدلی (بر مبنای مقیاس بارک) از بازنمایی شنیداری واکه.
- Terbeek, D. (1977) A cross-language multidimensional scaling study of vowel perception. *UCLA Working Papers in Phonetics*, 37, 1-271.
- یک مطالعه زبان‌گذر وسیع در زمینه درک واکه که نقطه شروعی است برای هر کس که بخواهد درک واکه را مطالعه کند.
- Traunmüller, H. (1981) Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America*, 69, 1465-75.
- مدل دیگری از بازنمایی شنیداری واکه.

تمرین

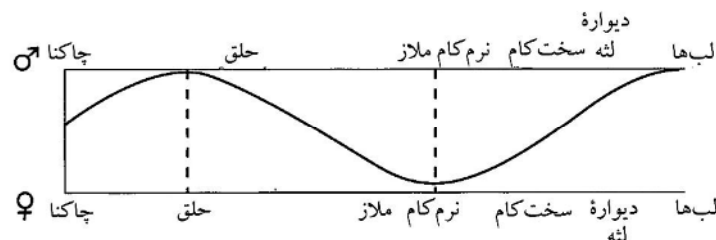
واژگان تخصصی

اصطلاحات زیر را تعریف کنید: نظریه منبع-فیلتر، سطح مقطع عرضی، مدل لوله‌ای، نوموگرام، بازخوان هلمهولتز، نظریه اختلال، پادگره، گره، پراکندگی تطبیقی، واکه‌های [افراشته] (برحسب F_1)، واکه‌های [افتاده] (برحسب F_1)، واکه‌های [پیشین] (برحسب F_2)، واکه‌های [پسین] (برحسب F_2).

پرسش

۱. با توجه به شکل‌های ۱-۶ و ۲-۶، فرکانس F_1 و F_2 وقتی طول حفره عقبی ۵ سانتی‌متر است، چقدر است؟ آیا F_1 یک بازخوان حفره جلویی است یا یک بازخوان حفره عقبی؟ فرکانس F_1 و F_2 (و وابستگی حفره‌ای آنها) وقتی طول حفره جلویی ۵ سانتی‌متر است، چه مقدار است؟
۲. برطبق معادله (۲-۶)، مساحت گرفتگی (A_c) در فرکانس F_1 نقش دارد. این ادعا را با محاسبه F_1 برای مقادیر $A_b = 3 \text{ cm}^2$, $l_b = 4 \text{ cm}$, $l_c = 2 \text{ cm}$ ، و فرض تغییر سطح مقطع گرفتگی از $A_c = 0.05 \text{ cm}^2$ تا $A_c = 0.2 \text{ cm}^2$ اثبات کنید. فرکانس F_1 وقتی $A_c = 0 \text{ cm}^2$ است، چقدر می‌باشد؟
۳. شکل ۱۲-۶ موج ایستاده F_3 به همراه محدوده‌های^۱ تولیدی برای گویشوران زن و مرد را به صورت کلی نشان می‌دهد. با استفاده از نظریه اختلال، تولیدهایی که پایین‌ترین مقدار F_3 را در [I] ایجاد می‌کنند، توصیف کنید. چگونه این «بهترین» تولیدها بین زنان و مردان متفاوت‌اند؟

1. landmark



شکل ۶-۱۲. موج ایستاده سومین سازه بازخوانی (در لوله‌ای که در یک انتها باز و در انتهای دیگر بسته است)، به همراه محدوده‌های تولیدی برای گویشوران زن و مرد. اقتباس از هگیوارا^۱ (۱۹۹۵، صفحه ۱۲). چاپ با کسب اجازه.

۴. فرکانس اولین و دومین سازه را در نقطه میانی واکه‌های مزاتک که در شکل ۶-۸ نشان داده شد، اندازه‌گیری کنید و آنها را با میانگین مقادیر در شکل ۶-۹ رسم نمایید. گویشوری که کلمات شکل ۶-۸ را تولید کرده، چه تفاوتی با میانگین دارد؟ آیا در این مقایسه، شاهی مبنی بر اینکه دستگاه گفتار وی کوتاه‌تر یا بلندتر از میانگین باشد، وجود دارد؟ آیا شاهی مبنی بر اینکه سرعت یا سبک گفتاری او با دیگران متفاوت است، وجود دارد؟

۵. فضای آکوستیکی یک واکه مرکب چگونه است؟ مقادیر سازه‌ای واکه مرکب شکل ۴-۹ را در ۴ یا ۵ نقطه با فواصل زمانی مساوی اندازه بگیرید و مقادیر را در یک فضای واکه‌ای آکوستیکی (متصل به هم با یک خط) رسم کنید. نظریه پراکندگی تطبیقی (و احتمالاً نظریه ذره‌ای) مبتنی بر این فرض هستند که مقوله‌های واکه‌ای از هم متمایز می‌شوند، زیرا موقعیت‌های متفاوتی در فضای واکه‌ای آکوستیکی دارند. چگونه واکه‌های مرکب نظریه واکه‌های «مرجح» را پیچیده می‌کنند؟

۶. در شکل ۶-۱۰ (ب)، قله کوچک بین F_0 و F_1 در [e]، [a] و [o] چیست؟ چرا این قله در این طیف‌ها ظاهر شده است؟ و چگونه این جنبه طیف‌های شنیداری می‌تواند پیداکردن F_1 در حلزون‌نگاشت را با مشکل مواجه کند؟

1. Hagiwara

فصل ۷

سایشی‌ها

نظریهٔ آکوستیکی تولید گفتار علاوه بر توصیف نظری مشخصه‌های آکوستیکی واکه‌ها، ویژگی‌های آکوستیکی همخوان‌های سایشی را نیز توصیف می‌کند. منبع نوفه (انرژی آکوستیکی) در سایشی‌ها، جریان هوای اغتشاشی^۱ است که با فرار هوا از یک انسداد تنگ در محلی در دستگاه گفتار تولید می‌شود. این نوفهٔ نامتناوب سپس به‌وسیلهٔ دستگاه گفتار فیلتر می‌شود. در این فصل، منبع نوفهٔ اغتشاشی و نحوهٔ عملکرد صافی دستگاه گفتار در سایشی‌ها توصیف می‌شود. سپس، برخی جنبه‌های ذره‌ای در نگاشت تولید به‌آکوستیک سایشی‌ها معرفی می‌شود و مقایسه‌ای نیز بین ویژگی‌های شنیداری با آکوستیکی سایشی‌ها صورت می‌گیرد. فصل با مبحثی در رابطه با درک سایشی‌ها پایان می‌پذیرد.

۷-۱. اغتشاش

وقتی درون یک نی می‌دمید، نوفهٔ نامتناوب سوت‌مانندی تولید می‌شود. این

1. turbulent airflow

نوفه پیامد آکوستیکی جابه‌جایی نامنظم مولکول‌های هواست که با خروج هوا از یک مجرای نسبتاً تنگ ایجاد می‌شود. وقتی هوای پرسرعت درون مجرا با هوای ساکن بیرون برخورد می‌کند، جریان هوا دچار آشفتگی می‌شود (ناحیه اصلی تولید نوفه درون ناحیه‌ای به طول هشت برابر شعاع، خارج از انتهای نی است). عوامل اصلی تعیین‌کننده اینکه آیا جریان هوا اغتشاشی است یا نه، عبارت‌اند از اندازه مجرا و سرعت حجمی جریان هوا (حجمی از هوا که از یک نقطه خاص در واحد زمان می‌گذرد). به عنوان مثال، اگر ۱۰۰ سانتی‌متر مکعب هوا در ثانیه از یک مجرا عبور کند، اگر سطح مجرا کمتر از ۱۰ میلی‌متر مربع باشد **جریان هوای اغتشاشی** تولید می‌شود، ولی اگر سطح مجرا ۲۰ میلی‌متر مربع باشد، جریان هوای اغتشاشی تولید نمی‌شود. از این رو، جریان هوای اغتشاشی خیلی آسان‌تر با یک نی باریک تولید می‌شود تا با یک نی گشاد.

آیرودینامیک^۱ بزرگراه‌ها و سایشی‌ها

رفتار اتومبیل‌ها در یک بزرگراه مثال خوبی برای بسیاری از ویژگی‌های آیرودینامیکی سایشی‌هاست.

سرعت ذره‌ای^۲، یعنی سرعت ذرات هوا مانند سرعتی که کیلومترشمار خودرو نشان می‌دهد. **سرعت حجمی**، یعنی تعداد ذرات هوایی که از یک نقطه معین می‌گذرند، مانند شمارش تعداد ماشین‌ها به وسیله دستگاه خودروشمار که تعداد ماشین‌هایی که در واحد زمان از یک مکان معین عبور می‌کنند، را نشان می‌دهد.

جریان لایه‌ای^۳ وقتی روی می‌دهد که همه خودروها در باند خود باقی می‌مانند و به آرامی در امتداد بزرگراه حرکت می‌کنند. **جریان اغتشاشی** وقتی روی می‌دهد که خودروها مدام باند خود را تغییر می‌دهند؛ در جریان اغتشاشی، علاوه بر حرکت به سمت جلو، برخی حرکت‌های اتفاقی از یک سو به سوی دیگر نیز وجود دارد.

←

1. aerodynamic

2. particle velocity

3. laminal flow

→
اغتشاش مجرای^۱ در بزرگراه وقتی روی می‌دهد که بزرگراه از دو باند به شش باند تعریض می‌شود. اتومبیل‌ها به سرعت از مسیر باریک خارج شده و با پهن شدن جاده تغییر باند می‌دهند. **اغتشاش مانعی^۲** مانند تغییر باندی است که به دلیل توقف یک خودرو در بزرگراه روی می‌دهد. خودروها باند خود را برای دورزدن مانع تغییر می‌دهند.

مقاومت آیرودینامیکی^۳، مقاومت یک مجرا در برابر جریان هواست که در تفاوت بین جاده‌های دوبانده و چهاربانده دیده می‌شود. مجراهای بزرگ‌تر معمولاً ظرفیت سرعت حجمی بالاتری نسبت به مجراهای باریک‌تر دارند. به عبارت دیگر، مجراهای بزرگ‌تر مقاومت کمتری در برابر جریان هوا دارند.

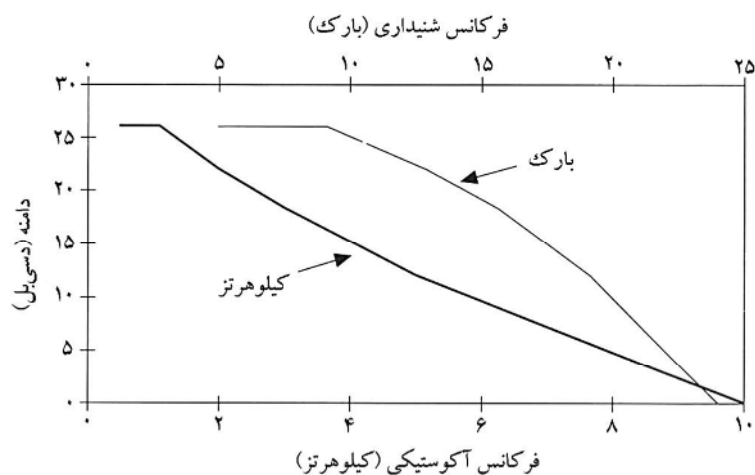
اگر سرعت حجمی ثابت باقی بماند، سرعت ذره‌ای با کم شدن پهنای مجرا افزایش پیدا می‌کند. وقتی تعداد باندها از چهار به دو کاهش پیدا می‌کند، سرعت خودروها درست قبل از این تغییر باند به شدت افت می‌کند، اما همین که خودروها وارد قسمت باریک می‌شوند، دوباره حرکت آنها شتاب می‌گیرد. برای آنکه همان تعداد اتومبیلی که از قسمت پهن عبور می‌کرد از قسمت باریک نیز عبور کند (سرعت حجمی یکسان)، اتومبیل‌ها باید در قسمت باریک تندتر حرکت کنند (سرعت ذره‌ای بالاتر).

حرکت مولکول‌های هوا در جریان هوای اغتشاشی، نامنظم است. به همین دلیل، امواج صوتی حاصل از جریان هوای اغتشاشی، مانند امواج نوفه سفید، تصادفی هستند. اما برخلاف طیف نوفه سفید، طیف **نوفه اغتشاشی** که معمولاً در سایشی‌ها دیده می‌شود، کاملاً مسطح نیست و همان‌طور که شکل ۷-۱ نیز نشان می‌دهد، دامنه طیف نوفه اغتشاشی، بالای ۱۰۰۰ هرتز با افزایش فرکانس به تدریج کاهش پیدا می‌کند.

1. channel turbulence

2. obstacle turbulence

3. aerodynamic impedance



شکل ۷-۱. طیف قدرت نوفه اغتشاشی (رجوع کنید به شیدل^۱، ۱۹۸۵). خط پررنگ، طیف را برحسب کیلوهرتز و خط کم‌رنگ، آن را برحسب مقیاس فرکانسی بارک نشان می‌دهد.

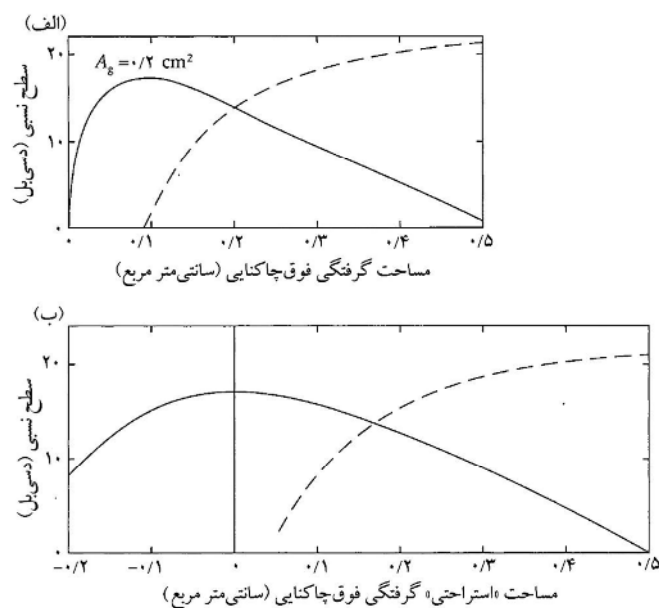
این نوع نوفه در سایشی‌ها و بسیاری از آواهای بی‌واک تولید می‌شود. به‌عنوان مثال، اغتشاش در [h] وقتی تولید می‌شود که هوای شش‌ها از چاکنای عبور می‌کند، زیرا منفذ بین تارهای صوتی نسبتاً تنگ و سرعت جریان هوا نسبتاً بالاست. (سرعت جریان هوا در دمش^۲ نسبتاً بالاست و آن را می‌توان با مقایسه مدت‌زمانی که می‌توانید تولید [h] را در مقایسه با [a] ادامه دهید، متوجه شد). از این رو، تفاوت اصلی بین [h] و واکه در کلمه heed این است که منبع صوتی در [h]، یک نوفه اغتشاشی نامتناوب است که در چاکنای تولید شده، در حالی که منبع صوتی در [i] یک موج متناوب مرکب (واک) است. هم واکه و هم [h] مقادیر سازه‌ای دارند که ناشی از صافی دستگاه گفتار است (هرچند نای در فیلتر آکوستیکی منبع صوت در [h]، بیش از واکه‌ها، مشارکت دارد). سایر سایشی‌ها نیز انرژی آکوستیکی در بازخوان‌های مشخصی در دستگاه گفتار دارند، اما قبل از آنکه به بحث درباره محل تولید سایشی‌ها بپردازیم، دو نکته دیگر در زمینه نوفه اغتشاشی لازم به ذکر است.

نخست اینکه، دامنه نوفه اغتشاشی بر حسب سرعت ملکول‌های هوا وقتی که از

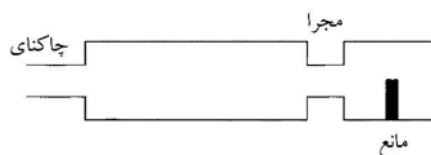
1. Shadle

2. aspiration

یک مجرا می‌گذرند، تعیین می‌شود (منظور سرعت ذره‌ای است، به مستطیل «آیرو دینامیک بزرگراه‌ها و سایشی‌ها رجوع کنید.»). هر قدر ملکول‌های هوا سریع‌تر حرکت کنند، صدا هم بلندتر است. از آنجا که سرعت ذره‌ای به مساحت مجرا بستگی دارد، می‌توان گفت برای یک مقدار معین جریان هوای در حال خروج از دهان (سرعت حجمی)، هر قدر مجرا تنگ‌تر باشد، نوفه اغتشاشی بلندتر است. این رابطه در شکل ۷-۲ (الف) نشان داده شده است. در این شکل دامنه نوفه سایش بر روی محور عمودی و مساحت گرفتگی بر روی محور افقی نشان داده شده است. شکل ۷-۲ (ب) نیز همین رابطه را نشان می‌دهد، اما با این فرض که فشار هوا کمی گرفتگی را باز می‌کند.



شکل ۷-۲. دامنه نسبی، تابعی است از مساحت یک گرفتگی اغتشاش‌زا در دستگاه گفتار. خطوط پیوسته، دامنه اغتشاش تولیدشده در محل گرفتگی و خطوط شکسته، دامنه اغتشاش چاکنایی را و هریک را به صورت تابعی از سایش گرفتگی نشان می‌دهند. شکل (الف) نتایج یک شبیه‌سازی را نشان می‌دهد، با این فرض که سطوح دستگاه گفتار غیرقابل تغییرند. شکل (ب) همان نتایج را نشان می‌دهد، اما با این فرض که مساحت گرفتگی [مساحت استراحتی] ^۱ ممکن است به وسیله فشار هوای پشت گرفتگی گسترده شود. A_g مساحت چاکناست. (اقتباس از استیونس (۱۹۸۹، ص. ۲۲).



شکل ۳-۷. مدل لوله‌ای یک سایشی منبع مانع. در این مدل، مجرای مولد جریان تند هوا و یک مانع در مسیر جریان دیده می‌شود. این مدل اغلب مربوط به همخوان‌های صغیری ^۱ [s] و [ʃ] است.

دوم اینکه، نوفه اغتشاشی علاوه بر آنکه با خروج جریان تندی از هوا از یک مجرای تنگ تولید می‌شود، با برخورد جریان تندی از هوا با مانعی در مسیر جریان نیز تولید می‌شود. شکل ۳-۷ نمونه‌ای از چنین شکل‌بندی را نشان می‌دهد. وجود یک مانع سبب افزایش دامنه نوفه اغتشاشی می‌شود (مقایسه کنید تعویض باندی که به دلیل تعریض بزرگراه روی می‌دهد با تعویض باندی که ناشی از مسدود شدن جاده به وسیله یک اتومبیل است). نوفه نه در فاصله مشخصی دورتر از گرفتگی، بلکه در مانع تولید می‌شود. می‌توان گفت که تقریباً همه نوفه‌های سایشی مستلزم اغتشاشی هستند که با برخورد جریانی از هوا با یک مانع تولید می‌شود. در [s] و [ʃ] به ترتیب دندان‌های بالا و پایین به عنوان موانع تولید اغتشاش عمل می‌کنند. اگر دقت کنید می‌بینید که حتی دامنه [f] وقتی شما لب فوقانی را (با دست) بالا نگه می‌دارید، کاهش پیدا می‌کند. این نشان می‌دهد که لب بالا به عنوان یک مانع در تولید سایشی لبی-دندانی نقش ایفا می‌کند. (این ممکن است دلیل یافته‌های مدیسون ۱۹۸۴ باشد که سایشی‌های لبی-دندانی در زبان‌های دنیا رایج‌تر از سایشی‌های دولبی هستند، زیرا اغتشاش مانع بلندتر از اغتشاش مجراست.)

شیدل (۱۹۹۱) بین سایشی‌های منبع «دیوار»^۲ که در آنها جریان تند هوا به یک دیواره در دستگاه گفتار برخورد می‌کند، مثل سایشی نرم‌کامی [x]، با سایشی‌های منبع «مانع»^۳ که در آنها جریان تند هوا با دندان‌های جلو برخورد می‌کند، تفاوت قائل می‌شود. به این دو دسته می‌توانیم سایشی‌های منبع «لب» مثل [v] و [f] را بیفزاییم که در آنها جریان هوا به لب بالا برخورد می‌کند. همه این انواع مختلف سایشی مستلزم

1. sibilant

2. wall source

3. obstacle source

یک مانع هستند. سایشی‌های بی‌مانع فقط در لب‌ها و شاید در چاکنای تولید می‌شوند. توجه کنید که چقدر تولید نوفه‌های سایشی دولبی و چاکنایی با صدای بلند دشوار است. در واقع، در برخی زبان‌ها، به‌عنوان مثال ناواهو^۱، واج /h/ با واج‌گونه‌های سایشی دهانی [x] یا [ç] تولید می‌شود. من این را در گفتار دقیق و محتاطانه در انگلیسی آمریکایی نیز دیده‌ام.

تفاوت بین سایشی‌های منبع «دیوار» با منبع «مانع» ناشی از جهت مانع نسبت به جهت جریان هواست. به‌عنوان مثال، در یک سایشی منبع «دیوار»، مانع تقریباً موازی با جریان هواست، در حالی که در یک سایشی منبع «مانع»، مانع تقریباً در زاویه عمود نسبت به جریان هوا قرار دارد. کتفورد^۲ (۱۹۷۷) اظهار می‌کند که وجود یک زاویه تند بین جهت مانع و جریان هوا سبب تولید جریان‌های حلقوی^۳ در جریان هوا می‌شود و از آنجا که این جریان‌ها متناوب هستند، سبب ایجاد یک موج متناوب مرکب با زیرویمی معینی می‌شوند (همان‌طور که در یک لوله آرگ روی می‌دهد). وی اشاره می‌کند که طیف‌های [s] تولیدشده توسط افراد با و بدون دندان مصنوعی (صفحه ۱۵۵) شاهی است بر اینکه جریان هوا، پشت دندان‌ها، تولید جریان‌های حلقوی متناوبی را می‌کند که منجر به ورود مؤلفه‌هایی با فرکانس بالا در طیف [s] می‌شود. شیدل (۱۹۹۱) مکان مانع را تغییر می‌دهد و تغییرات طیفی نسبتاً عظیمی مربوط به ویژگی‌های فیلتر حفره جلویی، هم در جلو و هم در عقب منبع نوفه نامتناوب مشاهده می‌کند.

سایشی‌های واک‌دار به‌نسبت در زبان‌های دنیا کمتر رایج‌اند. این سایشی‌ها دستخوش تغییرات آوایی گوناگونی می‌شوند و به‌طرز شگفت‌آوری تولید آنها دشوار است. دشواری تولید که ممکن است عامل اصلی الگوهای جهانی و واجی این همخوان‌ها باشد، ناشی از این است که تولید نوفه اغتشاشی ویژه سایشی‌ها، نیازمند سرعت حجمی بالاست، اما ارتعاش تارهای صوتی در برابر جریان‌یافتن هوا در دستگاه

۱. Navajo. یکی از زبان‌های بومی آمریکا. (مترجم)

گفتار مقاومت ایجاد می‌کند. برای مثال، اگر یک ورق کاغذ را روی صورت خود بگیرید، خیلی آسان‌تر می‌توانید آن را با سایشی لبی-دندانی بی‌واک [f] پایین بیندازید تا با سایشی لبی-دندانی واک‌دار [v]، زیرا در طول واک، تارهای صوتی به همان میزانی که باز هستند، بسته (یا تقریباً بسته) نیز هستند. بنابراین، با فرض مقدار مساوی فشار هوای تولیدشده به وسیله شش‌ها (فشار زیر حنجره‌ای)، سرعت حجمی در طی واک خیلی کمتر است نسبت به وقتی که چاکنای باز نگه داشته می‌شود. از آنجا که میزان مشخصی جریان هوا برای تولید اغتشاش لازم است، سایشی‌های واک‌دار ممکن است سایش خود را از دست بدهند و تبدیل به غلت^۱ شوند. توجه کنید که این تغییر آوایی ضرورتاً مستلزم تغییر در میزان گرفتگی دستگاه گفتار نیست؛ شما می‌توانید یک سایشی بی‌واک یا یک غلت واک‌دار با میزان گرفتگی یکسانی در دستگاه گفتار تولید کنید. مک‌دانو^۲ (۱۹۹۳) اظهار می‌کند که این موضوع، علت اصلی تناوب آوایی بین ناسوده‌های کناری^۳ واک‌دار با سایشی‌های کناری^۴ بی‌واک است. این تمایز واجی در برخی زبان‌ها ناشی از واک است که به دلیل پهنای مجرای کناری^۵ منجر به تولید سایش در همخوان بی‌واک و حذف سایش در همخوان واک‌دار می‌شود.

۲-۷. محل تولید سایشی‌ها

اگر شروع به تولید یک سری سایشی کنید، بدین صورت که با ایجاد یک گرفتگی در حلق، شروع به تولید سایش کنید، سپس گرفتگی را به سمت جلو تا دیواره لثه حرکت دهید، ممکن است بتوانید تغییری در «زیروبمی» سایش بشنوید. (تحلیل‌های این بخش برگرفته از هینز و استیونس^۶ ۱۹۶۱ است). تغییر فرکانسی که در این سایشی‌ها می‌شنوید، ناشی از تغییراتی است که در صافی دستگاه گفتار روی می‌دهد؛ به‌ویژه تغییراتی که در طول حفره جلویی دستگاه گفتار رخ می‌دهد. حفره جلویی در مدل لوله‌ای شکل ۷-۳، از مجرای تنگ یعنی نقطه گرفتگی تا لب‌ها گسترش پیدا می‌کند و مانع را نیز شامل می‌شود.

1. glide

4. lateral fricatives

2. McDonough

5. lateral fricatives

3. lateral approximant

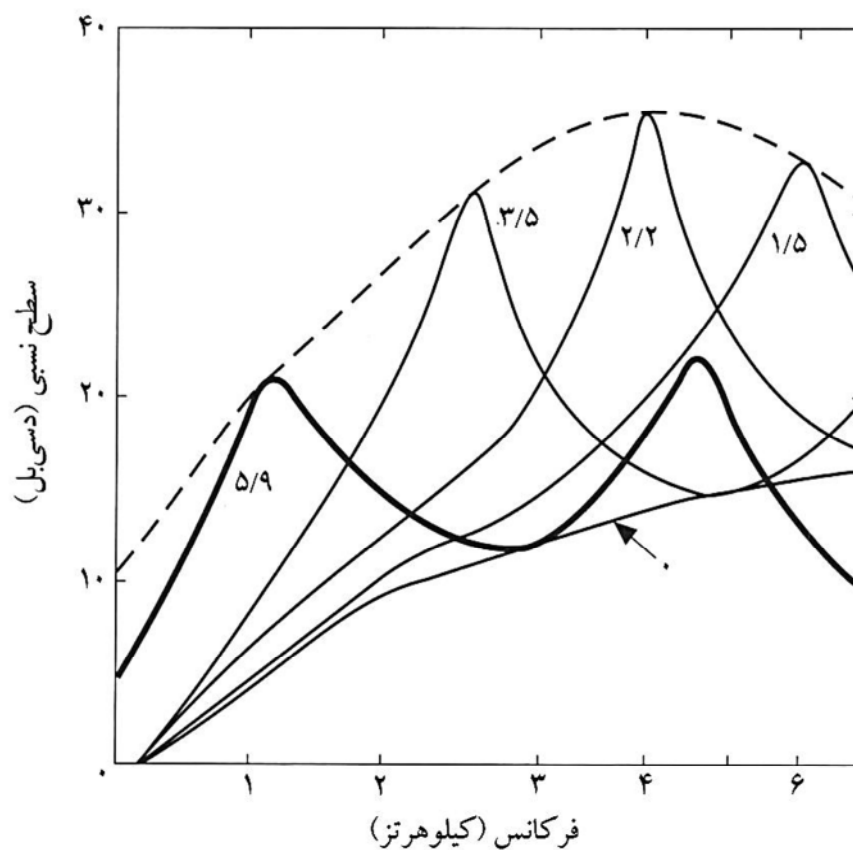
6. Heinz and Stevens

از آنجا که منبع صوت در حفره جلویی قرار دارد و از آنجا که جفت‌شدگی آکوستیکی بین حفره‌های جلو و عقب وقتی که دستگاه گفتار به شدت تنگ شده، ضعیف است، در نتیجه، فیلتر آکوستیکی دستگاه گفتار در سایشی‌ها در درجه اول به وسیله فرکانس‌های بازخوانی حفره جلویی صورت می‌گیرد (که می‌توانید آنها را با همان فرمولی که در فصل ۶ برای محاسبه بازخوان‌های واکه $F_n = (2n - 1)c/4L$ به کار رفت، به دست آورید). هنگامی که گرفتگی سایشی در حلق قرار دارد، طول حفره جلویی بلند است و در نتیجه، نسبت به وقتی که گرفتگی سایشی اندکی جلوتر در دهان قرار دارد، همخوان سایشی فرکانس‌های بازخوانی پایین‌تری دارد.

جفت‌شدگی آکوستیکی

گرفتگی‌ها تقریباً به همان صورتی که می‌توانند سبب مقاومت آیرودینامیکی شوند، می‌توانند سبب مقاومت آکوستیکی^۱ شوند، اما این دو نوع مقاومت با یکدیگر فرق دارند. به عنوان مثال، وقتی دهان و بینی خود را می‌بندید و یک صدای واک‌دار کوتاه مثل واک‌ی که طی یک انسدادی دولبی واک‌دار تولید می‌شود، تولید می‌کنید، هیچ هوایی از دستگاه گفتار شما خارج نمی‌شود، اما صدا خارج می‌شود. دلیل آن، این است که امواج صوتی تولیدشده در چاکنای می‌توانند در گونه‌ها به ارتعاش درآیند، همان‌طور که صدای استریوی همسایه در دیوار شما به ارتعاش درمی‌آید (البته فیلتر آکوستیکی اجسام جامد فقط نوفه‌هایی با فرکانس پایین را عبور می‌دهد). **جفت‌شدگی آکوستیکی** به مقاومت آکوستیکی، یعنی درجه انتقال صوت در برخی موانع آکوستیکی بستگی دارد. در سایشی‌ها، حفره‌های جلو و عقب به خوبی جفت نمی‌شوند و مقاومت آکوستیکی در محل گرفتگی نسبتاً زیاد است. از این رو، هرچند که هوا در گرفتگی جریان پیدا می‌کند، صدای نسبتاً کمی از حفره عقبی و یا به درون آن منتقل می‌شود. در نتیجه، فرکانس‌های بازخوانی حفره عقبی تأثیر نسبتاً کمی بر طیف‌های سایشی دارند.

1. acoustic impedance



شکل ۴-۷. طیف‌های سایشی که به وسیله یک مدل لوله‌ای شبیه به مدل لوله‌ای شکل ۳-۷ تولید شده‌اند. طیف‌ها در پنج مدل با تغییر طول حفره جلویی که برای هر طیف (برحسب سانتی‌متر) مشخص شده است، تولید شده‌اند. اقتباس از استیونس (۱۹۸۹، ص. ۲۵).

شکل ۴-۷، طیف‌های سایشی تولیدشده در یک مدل دولوله‌ای را نشان می‌دهد. عدد کنار هر طیف بیانگر طول حفره جلویی (برحسب سانتی‌متر) است. هر قدر طول حفره جلویی کوتاه‌تر، فرکانس پایین‌ترین قله طیفی بالاتر است (به جز طیفی که حفره جلو ندارد و با عدد صفر مشخص شده است). همچنین، توجه کنید که وقتی طول حفره جلویی ۲/۲ سانتی‌متر است، قله طیفی بالاترین دامنه را دارد. استیونس (۱۹۸۹) کمتر بودن دامنه را وقتی که طول حفره جلویی ۱/۵ سانتی‌متر است، ناشی از افزایش

میرایی آکوستیکی در انتشار صداهای دارای فرکانس بالا از لب‌ها می‌داند (دامنه نسبی طیف، ۶ دسی‌بل در هر اکتاو با انتشار صوت از لب‌ها افزایش پیدا می‌کند). وی همچنین اظهار می‌کند که کاهش دامنه سایشی‌هایی که گرفتگی آنها در دهان عقب‌تر است، ناشی از «کم‌بودن درجه جفت‌شدگی منبع با بازخوان حفره جلویی» است. من نمی‌دانم این، چرا باید روی دهد. وقتی برای اولین بار این شکل را دیدم، به اشتباه تصور کردم که وی میرایی آکوستیکی را مدل‌سازی کرده است که وقتی دیواره‌های لوله آکوستیکی نرم هستند، روی می‌دهد. بنابراین انتظار می‌رود لوله‌هایی که فضای دیواری بیشتری دارند (لوله‌هایی با حفره جلویی بلندتر)، میرایی آکوستیکی بزرگ‌تر و در نتیجه بازخوان‌هایی با پهنای باند وسیع‌تر و قله‌هایی با دامنه کمتر داشته باشند (برای مطالب بیشتر در زمینه پهنای باند به فصل ۹ رجوع کنید).

هرچند که موقعیت انرژی آکوستیکی در طیف، بیشتر آنچه را که درباره محل تولید سایشی‌ها باید بدانیم، به دست می‌دهد، اما حضور یا عدم حضور قله‌های بازخوانی دستگاه گفتار نیز می‌تواند یک سرنخ آکوستیکی برای محل تولید این همخوان‌ها باشد. بعضی سایشی‌ها مثل [f] که خیلی جلوی دهان تولید می‌شوند، ممکن است هیچ فیلتری در دستگاه گفتار نداشته باشند (نگاه کنید به طیف همخوان سایشی با طول حفره جلویی صفر در شکل ۷-۴). برخلاف واکه‌ها، مکان منبع نوفه در سایشی‌ها تغییر می‌کند. در سایشی‌های حلقی [h] و [ʔ] منبع انرژی آکوستیکی در حلق است و طول بازخوان دستگاه گفتار نسبتاً بلند است، اما در سایشی‌های لبی مثلاً [f] و [v] که با ایجاد اغتشاش در لب‌ها تولید می‌شوند، تقریباً هیچ عضوی در جلوی گرفتگی برای فیلتر صدا وجود ندارد. بنابراین، طیف سایشی‌های لبی قله ندارد و انرژی در محدوده فرکانسی بزرگی گسترده است (مثل طیف اغتشاش شکل ۷-۱). یاکوبسن و همکاران (۱۹۵۲) اصطلاح [فشرده^۱] را برای طیف‌هایی که فقط یک قله برجسته بازخوانی دارند، به کار بردند و اصطلاح [پراکنده^۲] را برای طیف‌هایی که

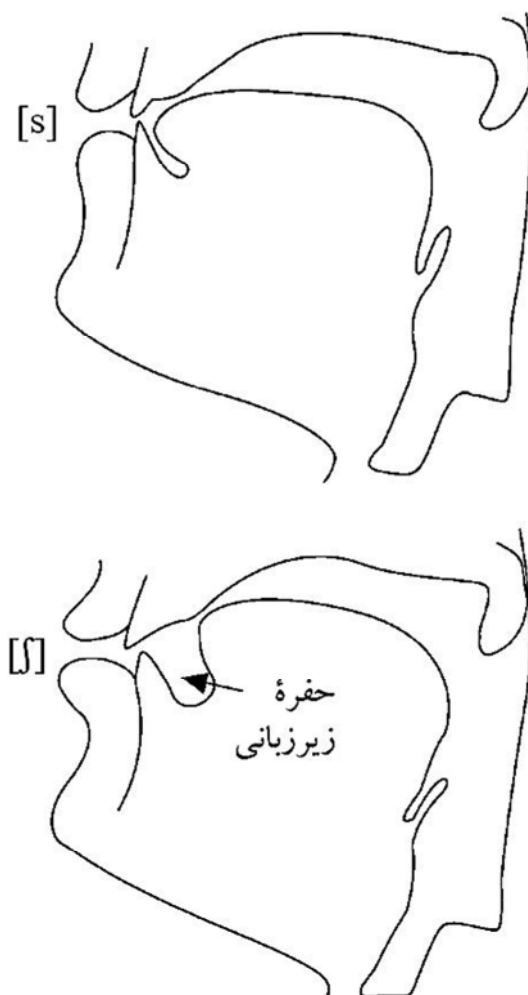
1. compact

2. diffuse

بدون فیلتر یا با فیلتر کم در دستگاه گفتار تولید می‌شوند و یا برعکس با یک حفره جلویی بلند و بنابراین، با چندین قله طیفی تولید می‌شوند، به کار بردند.

گردی لب‌ها همان تأثیر آکوستیکی را که در تولید واکه‌ها دارد، در تولید سایشی‌ها نیز دارد: فرکانس‌های سازه‌ای را پایین می‌آورد. به‌عنوان مثال [s] انگلیسی آمریکایی در کلمه see گرد نیست اما در sue تحت تأثیر هم‌تولیدی با [u] گرد است. شاید شما بتوانید تفاوت بین این دو آوای [s] را بشنوید. جالب است که دیگر سایشی تیز^۱ انگلیسی آمریکایی [ʃ] نیز (که در آن دندان‌های بالا به‌عنوان یک مانع عمل می‌کنند) علاوه بر آنکه محل تولید متفاوتی دارد، معمولاً با گردی لب‌ها تولید می‌شود. این تولید اضافی اختلاف آکوستیکی بین [s] و [ʃ] را افزایش می‌دهد.

تولید [ʃ] در بسیاری از زبان‌ها تحت تأثیر وجود یک حفره زیرزبانی نیز هست (شکل ۷-۵). با یک بررسی ساده در بین گویشوران بومی زبان انگلیسی در اطراف آزمایشگاه زبان‌شناسی دانشگاه ایالتی اوهایو دریافتیم که هفت نفر از هشت گویشور، [ʃ] را با یک حفره زیرزبانی تولید می‌کنند. (شما می‌توانید این را با یک خلال دندان آزمایش کنید. طی تولید ممتد سایشی، خلال دندانی بین دندان‌هایتان وارد کنید و به زبان‌تان ضربه بزنید. اگر بتوانید به زیر زبان ضربه بزنید، شاهدی است بر وجود یک حفره زیرزبانی در هنگام تولید). این فضای زیرزبانی به‌طور مؤثری به طول حفره جلویی دستگاه گفتار می‌افزاید و در نتیجه، فرکانس بازخوانی آن را پایین می‌آورد. از این رو، هرچند که [s] و [ʃ] محل‌های تولید تقریباً یکسانی دارند، خصوصیات آکوستیکی آنها به‌وضوح متفاوت است، زیرا [ʃ] یک حفره زیرزبانی دارد، اما [s] یا یک حفره زیرزبانی خیلی کوچک دارد ([s] با نوک زبان بالا) و یا اصلاً هیچ حفره‌ای ندارد ([s] با نوک زبان پایین).



شکل ۷-۵. ردیابی اشعه^۱ X از تولید آواهای [s] و [z] در زبان انگلیسی، حفرة زیرزبانی در [z] را نشان می‌دهد. (اقتباس از استراکا^۲ ۱۹۶۵، ۳۸).

۷-۳. نظریه ذره‌ای و سایشی‌ها

دو نوموگرامی که رابطه بین پارامترهای تولیدی سایشی با خصوصیات آکوستیکی آنها را نشان می‌دهند، نواحی ذره‌ای (نواحی پایدار آکوستیکی) را نشان

1. x-ray tracing 2. Straka

می‌دهند. اولی که رابطه بین میزان گرفتگی با دامنه نسبی سایشی را نشان می‌دهد، به صورت شکل ۷-۲ بازتولید شد. این شکل، مبنای ذره‌ای تمایزهای «نحوه تولید» یا «انسداد» را نشان می‌دهد. در محدوده‌ای از سایز گرفتگی، درجه گرفتگی هر اندازه باشد، دامنه نوفه سایش اساساً تغییر نمی‌کند؛ به‌ویژه وقتی که تأثیر فشار هوای دهانی بر روی اندازه مجرا نیز در نظر گرفته شود (شکل ۷-۲ (ب) را که در آن، نواحی پایدار از ۰/۱- سانتی متر مربع تا ۰/۱ سانتی متر مربع وسعت دارد، ببینید). این یافته‌ها نشان می‌دهد که سایش یکی از صرفاً چند برون‌داد آکوستیکی طبیعی دستگاه گفتار است؛ سایر برون‌دادهای آکوستیکی عبارت‌اند از سکوت در همخوان‌های ناپیوسته^۱ و جریان هوای لایه‌ای در آواهای رسا^۲. به عبارت دیگر، درجات مختلف انسداد که در نظریه آوایی و واجی به آنها اشاره می‌کنیم، طبقات طبیعی آکوستیکی هستند، زیرا درجات مختلف گرفتگی منجر به تولید انواع پایدار برون‌داد آکوستیکی می‌شود. (اگر طبقه سایشی‌ها را به صورت آوایی که با جریان هوای اغتشاشی تولید می‌شوند، تعریف کنیم، در این صورت «سایشی‌های» چاکنایی [h] و [f] سایشی محسوب می‌شوند، اما این آواها برخلاف سایر سایشی‌ها، ناهمخوانی^۳ هستند، به این معنا که در آنها لوله‌های جلو و عقب جفت می‌شوند (یعنی فضای بین سازه‌ها شبه‌واکه‌ای است)، در حالی که در آواهای همخوانی، جفت‌شدگی حفره‌های جلو-عقب روی نمی‌دهد و از این رو، فضای بین سازه‌ها پهن‌تر است و در مرز بین واحدهای همخوانی و ناهمخوانی تغییرات ناگهانی در تعداد سازه‌ها مشاهده می‌شود).

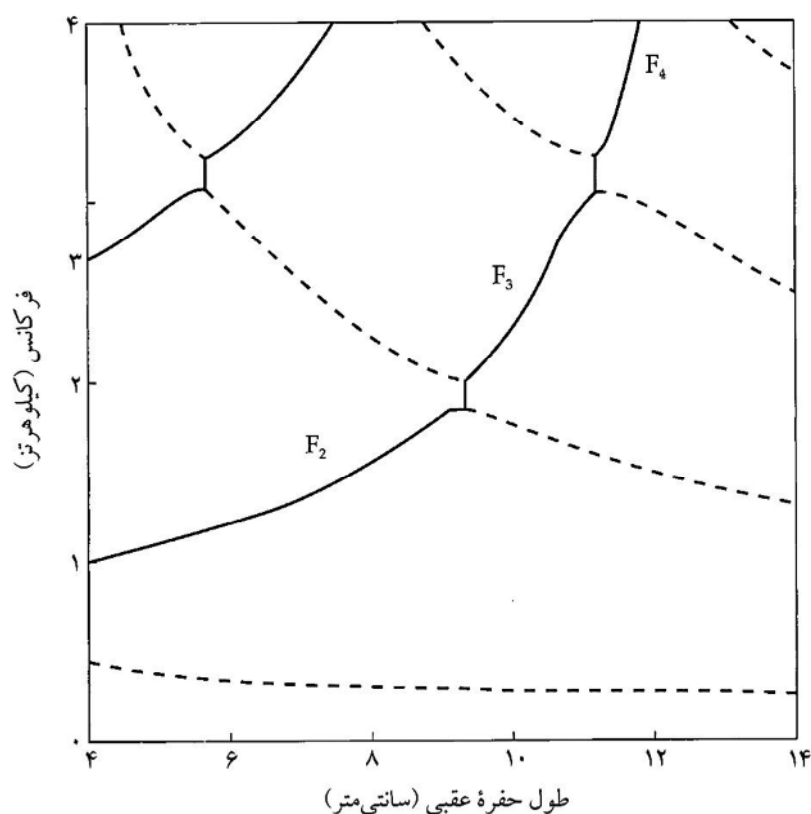
نوموگرام دوم در شکل ۷-۶ نشان داده شده است. این شکل، پیامد جفت‌شدگی لوله‌های جلو و عقب در یک مدل دولوله‌ای را نشان می‌دهد. همان‌طور که قبلاً نیز گفته شد، با کاهش طول حفره جلویی، فرکانس قله بازخوانی در سایشی افزایش پیدا می‌کند. هرچند که بازخوان‌های حفره عقبی در همخوان سایشی منتشرشده ظاهر نمی‌شوند یا صرفاً ضعیف ظاهر می‌شوند، اما جفت‌شدگی حفره‌های جلو و عقب بر

1. noncontinuant

2. sonorant

3. nonconsonantal

نوموگرام تأثیر می‌گذارد. به ناپیوستگی‌هایی توجه کنید که در فرکانس‌های بازخوانی حفره جلو روی می‌دهد، وقتی که فرکانس‌های بازخوانی حفره‌های عقب و جلو با هم تلاقی می‌کنند.



شکل ۶-۷. تأثیر جفت‌شدگی حفره‌های جلو و عقب دستگاه گفتار بر ویژگی‌های بازخوانی سایشی‌ها. فرکانس‌های بازخوانی حفره جلو با خط پیوسته و فرکانس‌های بازخوانی حفره عقب با خط چین نشان داده شده است. پایین‌ترین خط چین، بازخوان «هلمهولتز» در حفره عقبی را نشان می‌دهد. اقتباس از استیونس (۱۹۸۹، ص. ۲۶).

وقتی محل گرفتگی در عقب دستگاه گفتار است (یعنی وقتی طول حفره جلویی بلند و طول حفره عقبی کوتاه است، همان‌گونه که در قسمت چپ نوموگرام نشان داده شده)، پایین‌ترین فرکانس بازخوانی سایشی مطابق با دومین سازه واکه

مجاور است. از این رو، در [xa] (x) سایشی نرم‌کامی بی‌واک)، پایین‌ترین قله بازخوانی سایشی برابر با فرکانس F_2 در آغاز واکه بعدی است. در گرفتگی‌های کمی جلوتر در دستگاه گفتار، وقتی بازخوان حفره جلویی با بازخوان حفره عقبی تلاقی می‌کند، وابستگی سازه‌ای قله بازخوانی سایشی تغییر می‌کند. در نتیجه، فرکانس قله بازخوانی [ç] (سایشی سخت‌کامی بی‌واک) مطابق با فرکانس F_3 واکه بعدی است نه F_2 . آن از آنجا که فرکانس‌های بازخوانی دو لوله جفت‌شده هرگز دقیقاً یکسان نیست، در هر تلاقی بازخوان‌های حفره جلو و عقب (در هر تغییر وابستگی سازه‌ای قله بازخوانی همخوان سایشی) پرشی در نمودگرام دیده می‌شود. بر اساس نظر استیونس (۱۹۸۹)، چنین ناپیوستگی‌هایی مرز بین محل‌های تولید را که در سایشی‌ها، نگاشت تولید به آکوستیک نسبتاً پایداری دارند، مشخص می‌کنند.

۴-۷. طیف‌های شنیداری سایشی

شکل ۷-۷ طیف‌های آکوستیکی (ال‌پی‌سی) شش سایشی بی‌واک زبان عربی مصر را نشان می‌دهد. این طیف‌ها از کلمات زیر که توسط یک گویشور زن تولید شده‌اند، گرفته شده است:

[fæ:t]	گذراندن
[sæ:b]	او (مرد) رفت
[fæ:b]	خاکستری‌شدن، پیرشدن
[χæ:l]	دایی
[hæ:l]	حال
[hæ:t]	به من بده

بسیاری از جنبه‌های آکوستیکی سایشی‌ها که در این فصل مورد بحث قرار گرفت، در این طیف‌ها قابل مشاهده است. نگاه کنید، طیف [f] شباهتی به سایر طیف‌ها ندارد؛ از این لحاظ که هیچ برجستگی طیفی ندارد. انرژی تمایل به پخش‌شدن در سرتاسر

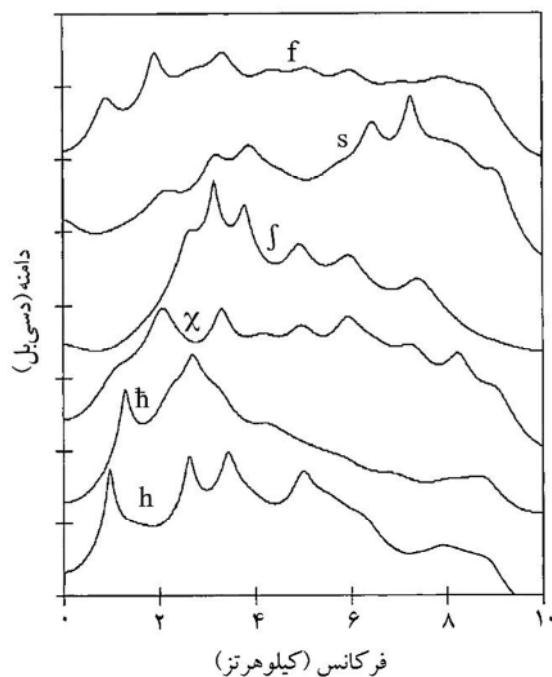
طیف دارد. سایر طیف‌ها برحسب موقعیت فرکانسی قله انرژی متفاوت‌اند؛ [s] بالاترین تمرکز فرکانسی و [h] تمرکز فرکانسی پایین‌تری را نسبت به سایر سایشی‌ها نشان می‌دهند. این یافته‌ها مطابق با پیش‌بینی قبلی ماست که فرکانس‌های بازخوانی حفره جلویی، تابع فیلتر دستگاه گفتار در سایشی‌ها را تعیین می‌کند و بنابراین، با جابه‌جایی محل تولید سایشی‌ها از جلو به عقب، قله انرژی طیفی در سایشی‌ها از فرکانس‌های بالاتر به فرکانس‌های پایین‌تر تغییر مکان می‌دهد. همچنین، توجه کنید که طیف [h] شبیه طیف یک واکه است هرچند که دامنه نسبی آن در سازه‌های بالاتر نسبت به آنچه که معمولاً در واکه‌ها می‌بینیم، بیشتر است.

به شکل ۷-۷ نگاه کنید. سایشی‌های [s] و [ʃ] تفاوت قابل ملاحظه‌ای با یکدیگر دارند. قله طیفی [ʃ] در حدود ۳/۵ کیلوهرتز است، در حالی که قله طیفی [s] نزدیک ۸ کیلوهرتز، هرچند که یک قله کوچک‌تر در حدود ۴ کیلوهرتز نیز دارد. پژوهشگران بسیاری معتقدند که اندازه‌گیری مشخصات آکوستیکی سایشی‌ها (به‌ویژه سایشی‌های تیغه‌ای) دشوار است، زیرا مانند این تولید [s]، ممکن است دارای چندین قله طیفی باشند و هریک از این قله‌ها ممکن است از تولیدی به تولید دیگری دارای بالاترین دامنه باشد. همچنین، در تعداد قله‌های طیفی سایشی‌ها نیز ممکن است تنوع بینافردی گسترده‌ای مشاهده شود. این مشاهدات منجر به گسترش روش‌های محاسبه مرکز ثقل^۱ برای توصیف طیف سایشی‌ها شد (فورست و همکاران^۲ ۱۹۸۸؛ جی‌سم^۳ ۱۹۷۹). بازنمایی شنیداری سایشی‌ها نیز جالب است. همان‌طور که در فصل ۴ گفته شد، دو تا از ویژگی‌های اصلی سیستم شنوایی انسان غیرخطی بودن مقیاس فرکانس شنیداری و پهنای نوارهای بحرانی شنیداری در فرکانس‌های بالاست. ترکیب این دو ویژگی به‌طور بنیادی بازنمایی شنیداری سایشی‌ها را در مقایسه با بازنمایی آکوستیکی آنها تغییر می‌دهد.

1. center of gravity

2. Forrest et al.

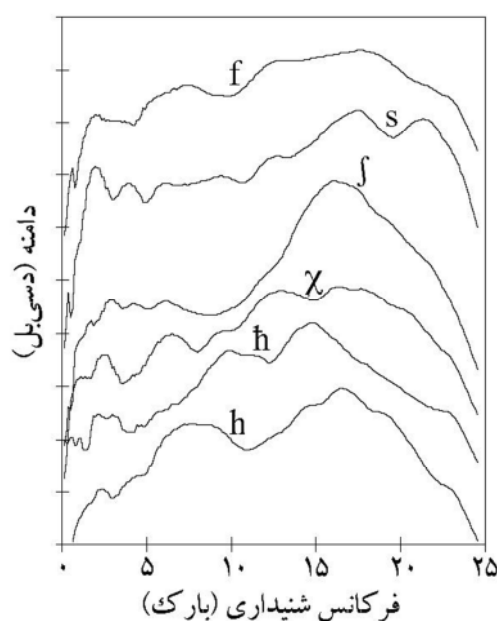
3. Jassem



شکل ۷-۷. طیف‌های آکوستیکی سایشی‌های عربی مصر که از کلمات فهرست‌شده در متن گرفته شده‌اند. طیف‌ها با تجزیه ال‌پی‌سی و با قراردادن پنجره شکل موج در میان نوفه سایشی به دست آمده‌اند و برای آنکه مقایسه شکل طیفی آنها راحت‌تر شود، (بدون در نظر گرفتن اختلاف دامنه کلی) کنار هم روی محور دامنه گذاشته شدند. فاصله بین علامت‌ها روی محور عمودی ۲۰ دسی‌بل است.

شکل ۷-۸. طیف‌های شنیداری همان همخوان‌های سایشی شکل ۷-۷ را نشان می‌دهد. این طیف‌ها با استفاده از همان پنجره‌های شکل موجی که برای تولید طیف‌های ال‌پی‌سی به کار رفت، محاسبه شده‌اند و بسیاری از همان ویژگی‌هایی را که در شکل ۷-۷ دیدیم، نشان می‌دهند. برای مثال، [f] یک طیف مسطح‌تر و پراکنده‌تر نسبت به سایر آواها دارد و در طیف سایر سایشی‌ها با جابه‌جایی محل گرفتگی از جلو به عقب، برجستگی‌های طیفی که به تدریج فرکانس پایین‌تر دارند، قابل مشاهده است. اما برخلاف طیف‌های آکوستیکی، به نظر می‌رسد که طیف‌های شنیداری این آواها به‌طور یکنواخت‌تری از یکدیگر فاصله دارد. همچنین، توجه کنید که اختلاف بین دو ناحیه بزرگ انرژی طیفی در [s] کاهش یافته است. معقولانه به

نظر می‌رسد اگر بگوییم که برخی از دشواری‌های مربوط به توصیف آکوستیکی سایشی‌ها در بازنمایی شنیداری آنها کاهش می‌یابد، اگر چه از بین نمی‌رود. در همین رابطه خوب است اضافه کنیم که تکنیک تجزیه طیفی که اغلب در سیستم‌های بازشناسی گفتار موجود در بازار به کار می‌رود، نوعی روش تجزیه شنیداری به نام رمزگذاری کپسترال^۱ در مقیاس فرکانسی مل^۲ (دیویس و مرملشتین^۳ ۱۹۸۰) را به کار می‌گیرد، زیرا مقیاس مل که یک مقیاس فرکانسی شنیداری است، ابهام در فرکانس‌های بالا را کاهش می‌دهد. در این فرکانس‌ها ممکن است تولید گویشوران به‌طور معناداری متفاوت، اما دارای کمترین تأثیر بر درک شنونده باشد.



شکل ۷-۸ طیف‌های شنیداری سایشی‌هایی که در شکل ۷-۷ نشان داده شد. طیف‌های شنیداری از همان پنجره‌های شکل موجی که برای محاسبه طیف‌های ال‌پی‌سی در شکل ۷-۷ به کار رفت، گرفته شده‌اند و مثل همان طیف‌ها، برای آنکه مقایسه شکل طیفی آنها راحت‌تر باشد، کنار هم روی محور دامنه قرار گرفتند. فاصله بین علامت‌ها روی محور عمودی ۲۰ دسی‌بل است.

1. cepstral coding

2. Mel frequency

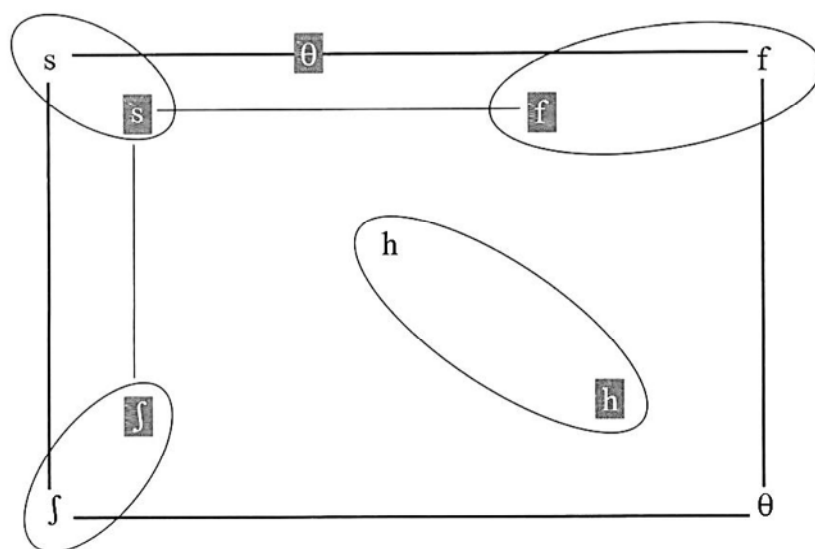
3. Davis and Mermelstein

بازنمایی شنیداری در شکل ۷-۸، علاوه بر نتایج عملی که برای اندازه‌گیری نوفه‌های سایشی و بازشناسی آنها در برنامه‌های بازشناسی گفتار دارد، نکته نظری جالبی را نیز مطرح می‌کند. با نگاهی به طیف‌های آکوستیکی شکل ۷-۷ (و دانستن اینکه گستره تمایزهای سایشی در زبان عربی معرف گستره تمایزهای سایشی یافت‌شده در زبان‌های دنیاست) عجیب به نظر می‌رسد که بیشتر ناحیه فرکانسی بالای ۵ کیلوهرتز برای تمایز نوفه‌های سایشی به کار نمی‌روند. چرا قله‌های طیفی چهار سایشی با هم زیر ۴ کیلوهرتز تجمع یافته‌اند، در حالی که فقط قله طیفی [s] بالای این مقدار است؟ به نظر می‌رسد که فقط نیمی از ناحیه فرکانسی زیر ۱۰ کیلوهرتز برای تمایز سیگنال‌های سایشی به کار می‌رود. این به‌ویژه عجیب به نظر می‌رسد که با توجه به این واقعیت (قبلاً ذکر شد) که [ʃ] معمولاً با یک حفره زیرزبانی و گردی لب‌ها که برای افزایش طول حفره جلویی دستگاه گفتار به کار می‌رود، تلفظ می‌شود و بنابراین، فرکانس قله طیفی مربوط به این آوا پایین می‌آید. این ویژگی عجیب طیف‌های آکوستیکی در طیف‌های شنیداری شکل ۷-۸ صادق نیست. در طیف‌های شنیداری، قله‌های طیفی سایشی در محدوده فرکانسی موجود تقریباً به‌طور یکنواخت‌تری از یکدیگر فاصله دارند و این یعنی که در زبان‌های دنیا، سایشی‌هایی که قله‌های طیفی آنها بین قله‌های [s] و [ʃ] باشد، رایج نیستند نه به این دلیل که تولید آنها مشکل است، بلکه به این دلیل که شنیدن آنها دشوار است.

۷-۵. ابعاد درکی سایشی‌ها

در شکل ۷-۹، آرایش سایشی‌های بی‌واک انگلیسی در فضای درکی نشان داده شده است. این شکل نتایج تحلیل دو مقیاس‌گذاری چندبعدی، یکی برای شنوندگان انگلیسی آمریکایی (حروف سیاه در زمینه سفید) و دیگری برای شنوندگان ژاپنی (حروف سفید در زمینه سیاه) را نشان می‌دهد. برای شنوندگان، محرک‌های یکسانی از زبان انگلیسی آمریکایی پخش شد و آنها را به‌صورت "f"، "th"، "s"، "sh" یا "h"

شناسایی کردند. من این شکل را با استفاده از ماتریس‌های ابهام منتشرشده توسط لمباچر و همکاران^۱ (۲۰۰۲) با روشی که در فصل ۵ توضیح داده شد، رسم کردم. ابتدا به فضای درکی شنوندگان انگلیسی آمریکایی نگاه کنید. این شنوندگان دچار ابهام خیلی کمی در شناسایی این آواها شدند (شناسایی در حدود ۹۴ درصد درست است). در این فضای دوبعدی پنج سایشی داریم که تقریباً تا حد ممکن متفرق شده‌اند: یک آوا در هر گوشه فضای درکی و یک آوا در میان آن. ملاحظه کنید که سایشی‌های صغیری یا تیز با هم در سمت چپ فضای درکی و سایشی‌های دهانی غیرصغیری [f] و [θ] با هم در سمت راست آن قرار دارند (برای شنوندگان انگلیسی آمریکایی). دو سایشی [s] و [f] با گرفتگی نسبتاً جلوتر در دستگاه گفتار، در بالای فضای درکی و دو سایشی نسبتاً عقب‌تر [θ] و [ɹ] در پایین آن قرار دارند.



شکل ۷-۹. تحلیل مقیاس‌گذاری چندبعدی داده‌های گزارش‌شده توسط لمباچر و همکاران^۲ (۲۰۰۱). فضای درکی شنوندگان انگلیسی آمریکایی با حروف سیاه در زمینه سفید و فضای درکی شنوندگان ژاپنی با حروف سفید در زمینه سیاه رسم شده است. نمونه‌ها توسط یک گویشور انگلیسی آمریکایی تولید شده‌اند.

1. Lambacher et al.

2. Lambacher et al.

همبسته آکوستیکی محتمل بُعد افقی، دامنه نوفه سایش است، [s] و [ʃ] نوفه سایش بلندتری نسبت به سایشی‌های دیگر دارند. بُعد عمودی ممکن است به شکل طیفی نوفه سایش مربوط باشد، [ʃ] و [θ] مؤلفه‌های فرکانس پایین بیشتری دارند. [h] طیف بسیار پراکنده‌ای دارد که می‌تواند توجیهی بر این باشد که چرا در میان فضای درکی قرار گرفته است.

اکنون به نتایج شنوندگان ژاپنی توجه کنید (حروف سفید). اولاً ملاحظه کنید که فضای ژاپنی در درون فضای انگلیسی آمریکایی جای گرفته است. این، یعنی که شنوندگان ژاپنی به‌طور کلی کمتر توانسته‌اند این آواهای انگلیسی آمریکایی را به درستی شناسایی کنند. دوم اینکه سه تا از سایشی‌ها در گوشه‌های فضای درکی قرار دارند همان‌طور که برای شنوندگان انگلیسی بودند. این آواها [f]، [s] و [ʃ] هستند. در مورد این آواها، فضای درکی ژاپنی و انگلیسی آمریکایی تقریباً شکل یکسانی دارد.

حالا به [θ] نگاه کنید. در فضای ژاپنی، [θ] به [s] نزدیک‌تر است تا به سایشی‌های دیگر، در حالی که برای شنوندگان انگلیسی آمریکایی هیچ دو سایشی این‌همه از یکدیگر فاصله ندارند. دو مسئله قابل توجه است. نخست اینکه، [θ] یک واج در زبان ژاپنی نیست و برخلاف سایر سایشی‌های مورد مطالعه، یک آوای کاملاً خارجی محسوب می‌شود. دوم اینکه، هرچند نوفه‌های سایشی [s] و [θ] کاملاً با یکدیگر فرق دارد، اما گذرهای سازه‌ای آنها (یک سرنخ مهم برای محل تولید) کاملاً شبیه یکدیگر است. شنوندگان ژاپنی ممکن است به این شباهت اتکا کنند و تفاوت نوفه‌های سایش را که برای شنوندگان انگلیسی آمریکایی بسیار معنادار به نظر می‌رسد، نادیده بگیرند.

آنچه در این مثال مختصر می‌بینیم این است که سایشی‌ها به‌لحاظ درکی بر روی ابعادی مرتب شده‌اند که در درجه اول آوایی و زبانی هستند. ابعاد زبانی در تضاد با ابعاد شنیداری یا آکوستیکی سایشی‌ها نیستند، اما نشان می‌دهند که چگونه یک سیستم زبانی حاوی تقابل‌ها ممکن است شنوندگان را به سمت برخی مشخصه‌های شنیداری هدایت و از توجه به برخی دیگر باز دارد.

منابعی برای مطالعه بیشتر

- Catford, J. C. (1977) *Fundamental Problems in Phonetics*, Bloomington: Indiana University Press.
- آشنایی عمومی با آواشناسی. کتابی که برخلاف بیشتر کتاب‌ها در این زمینه، به طور عمیقی به آیرودینامیک گفتار می‌پردازد. هرگاه دانشجویان از من می‌پرسند «در فصل تابستان قبل از ورود به دوره کارشناسی ارشد چه منابعی را می‌توان مطالعه نمود؟» من معمولاً این کتاب را پیشنهاد می‌کنم.
- Davis, S. and Mermelstein, P. (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP 28, 357-66.
- توصیف رمزگذاری کپسترال در مقیاس فرکانسی مل به‌عنوان یک بازنمایی طیفی.
- Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (1988) Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84, 1 15-23.
- مجموعه‌ای از پارامترها (برمبنای گشتاورهای آماری توزیع داده‌ها) را برای اندازه‌گیری مشخصه‌های سایشی پیشنهاد می‌کند.
- Heinz, J. M. and Stevens, K. N. (1961) On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America*, 33, 589-96.
- مطالعه کلاسیک آکوستیک سایشی‌ها.
- Jakobson, R., Fant, G., and Halle, M. (1952) *Preliminaries to Speech Analysis*, Cambridge, MA: MIT Press.
- کتابی کلاسیک در آواشناسی و واج‌شناسی که در آن مجموعه‌ای از مشخصه‌های تمایزدهنده واجی به‌لحاظ آکوستیکی توصیف می‌شوند.
- Jassem, W. (1979) Classification of fricative spectra using statistical discriminant functions. In B. Lindblom and S. Obman (eds.), *Frontiers of Speech Communication Research*, New York: Academic Press, 77-91.
- از صافی‌های نوارگذر متعددی برای محاسبه طیف سایشی‌ها استفاده می‌کند.
- Lambacher, S., Martens, W., Nelson, B., and Berman, J. (2001) Identification of English voiceless fricatives by Japanese listeners: The influence of vowel context on sensitivity and response bias. *Acoustic Science & Technology*, 22, 334-43.
- بررسی درک سایشی که در بخش ۷-۵ مطرح شد.
- Maddieson, I. (1984) *Patterns of Sounds*, Cambridge: Cambridge University Press.
- مطالعه‌ای در رده‌شناسی آوایی زبان‌های دنیا، که به‌عنوان منبع پیشنهادی در انتهای چندین فصل نیز به آن اشاره شده است. فصل سایشی‌های کتاب مذکور، برخی مباحث کلیدی این فصل را فراهم می‌کند.
- McDonough, J. (1993) The phonological representation of laterals. *UCLA Working Papers in Phonetics*, 83, 19-32.
- مطالعه رابطه بین آیرودینامیک سایشی‌ها با بازنمایی واجی آنها.
- Shadle, C. (1985) The acoustics of fricative consonants. *RLE Technical Report*, 506, MIT.
- مطالعه کامل نظری و مدل نوفه‌های سایشی. این منبع، شناختی بنیادین از آیرودینامیک و آکوستیک سایشی‌ها به دست می‌دهد.
- Shadle, C. (1991) The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, 19, 409-24.
- یک نسخه کوتاه‌تر و در دسترس‌تر از بخش‌هایی از شیدل (۱۹۸۵).
- Stevens, K. N. (1989) On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.
- «مقادیر ذره‌ای» سایشی را نیز مورد بحث قرار می‌دهد.

تمرین

واژگان تخصصی

این اصطلاحات را تعریف کنید: اغتشاش، جریان هوای اغتشاشی، سرعت ذره‌ای، سرعت حجمی، جریان لایه‌ای، جریان اغتشاشی، اغتشاش مجرای، اغتشاش مانعی، مقاومت آیرودینامیکی، نوفه اغتشاشی، سایشی‌های منبع دیوار، جفت‌شدگی آکوستیکی، [فشرده]، [پراکنده]، حفره زیرزبانی.

پرسش

۱. در برش‌های طولی دستگاه گفتار در شکل ۷-۵، نقاط دارای بالاترین سرعت ذره‌ای را (با فرض سرعت ذره‌ای ثابت در دهان) مشخص کنید. نقطه تولید نوفه را با فرض اینکه قطر منفذ متناسب با پهنای نشان داده شده در شکل است، نشان دهید.
۲. تحلیل آکوستیکی سایشی‌ها در این فصل بر پایه این فرض بود که جفت‌شدگی حفره‌های جلو و عقب در سایشی‌ها به لحاظ آکوستیکی روی نمی‌دهد، زیرا گرفتگی سایشی‌ها بسیار تنگ است. در چه لحظاتی طی تولید یک سایشی ممکن است این فرض نادرست باشد؟ و به نظر شما جفت‌شدگی آکوستیکی حفره‌های جلو و عقب چه تأثیری بر طیف سایشی‌ها دارد؟
۳. سایشی‌های واک‌دار و [h] را بر اساس نظر مطرح شده در این فصل که طبقات گرفتگی مثل انسدادی، سایشی و غیره به خوبی به لحاظ آکوستیکی توصیف می‌شوند، مورد بحث قرار دهید.
۴. شکل‌های ۶-۶ و ۶-۴ را با هم مقایسه کنید. در فصل ۶ گفتم که نواحی پایدار واکه‌ها در نقاط تلاقی پایین‌ترین بازخوان‌های جلو و عقب روی می‌دهد. اما در این فصل، این نقاط به عنوان نواحی دارای کمترین پایداری در نگاشت تولید به آکوستیک سایشی‌ها معرفی شده‌اند. برای مثال، واکه‌ای با طول حفره عقبی در حدود ۷ سانتی‌متر در یک ناحیه ناپایدار نوموگرام قرار دارد، اما یک سایشی با همین طول حفره عقبی در ناحیه پایدار آن محسوب می‌شود. چگونه می‌توان این یافته‌ها را با تلاش‌هایی که برای یافتن مشخصه‌های محل تولید یکسان، که هم برای واکه‌ها و هم برای همخوان‌ها کاربرد داشته باشد، ربط داد؟

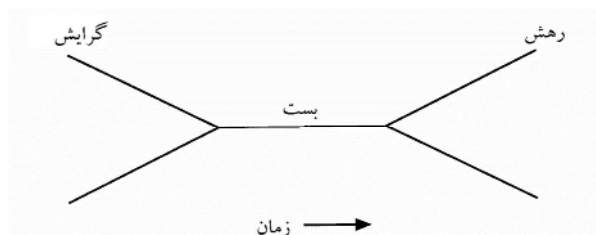
فصل ۸

انسدادی‌ها و انسایشی‌ها^۱

انسدادی‌ها و انسایشی‌ها ویژگی‌های آکوستیکی و تولیدی پیچیده‌تری نسبت به واکه‌ها یا سایشی‌ها دارند. در فصل‌های گذشته ویژگی‌های پویای واکه‌ها و سایشی‌ها نادیده انگاشته شد و این آواها به صورت رخداد‌های ساکن تلقی شدند. البته تا حدی منطقی به نظر می‌رسد که واکه‌ها و سایشی‌ها را به صورت اهداف آکوستیکی یا بسترهای تولیدی ساکن در نظر بگیریم. اما در مورد انسدادی‌ها و انسایشی‌ها چنین نیست. مهم‌ترین بستر تولیدی در طول یک همخوان انسدادی، بست کامل دستگاه گفتار است که پیامد آکوستیکی آن سکوت است (یا اگر تارهای صوتی ارتعاش کنند، واک ضعیف خواهد بود). اما زبان‌ها با به‌کارگیری محل‌های تولید مختلف و مکانیسم‌های مختلف تولید رهش انسدادی و نوفه‌های آن، انواع گسترده‌ای از همخوان‌های انسدادی بیشتر از آنچه این توصیف ساده پیش‌بینی می‌کند، به‌کار می‌برند. به‌همین دلیل تعیین سه مرحله در تولید انسدادی‌ها و انسایشی‌ها مطابق با سه بازه زمانی تولید آنها لازم است (شکل ۸-۱). هر یک از این سه مرحله را می‌توان بر اساس نظریه آکوستیکی تولید گفتار به صورت ترکیبی از یک منبع صوتی و یک تابع فیلتر دستگاه گفتار توصیف کرد. اولین مرحله عبارت است از حرکت

1. affricates

یک اندام تولیدی به سمت ایجاد یک بست انسدادی (مرحله گرایش)^۱، دومین مرحله خود بست و مرحله سوم رهش بست است.



شکل ۸-۱. در بازه زمانی تولید انسدادی یا انسایشی سه مرحله قابل تشخیص است. خطوط، اندام‌های تولیدی را نشان می‌دهند که در مرحله گرایش به سمت یکدیگر حرکت می‌کنند و در مرحله رهش از یکدیگر جدا می‌شوند.

۸-۱. توابع منبع در انسدادی‌ها و انسایشی‌ها

۸-۱-۱. انواع واک‌سازی

پیش از ورود به بحث توابع منبع در انسدادی‌ها و انسایشی‌ها، نگاهی به مشخصه‌های آکوستیکی انواع واک‌سازی می‌اندازیم، اگرچه ممکن است نوع واک‌سازی در واکه‌ها، سایشی‌ها و نیز انسدادی‌ها به طور تمایز دهنده‌ای با یکدیگر تفاوت داشته باشد. سه نوع اصلی واک در زبان‌ها تولید می‌شود که عبارت‌اند از واک عادی^۲، بازداشته و نفسی. واک عادی در همه زبان‌ها تولید می‌شود و برخی زبان‌ها آن را از یک یا هر دو نوع واک دیگر تمایز می‌دهند. هرچند که ارتعاش تارهای صوتی یک فرایند به‌غایت پیچیده است (برای مثال به بلس و آبز^۳ ۱۹۸۳ و لیور^۴ ۱۹۸۰ رجوع کنید). اما من فقط روی یک جنبه که به نظر می‌رسد به‌لحاظ آکوستیکی نیز مرتبط‌ترین مورد است، تمرکز می‌کنم: مدت زمانی که تارهای صوتی طی یک سیکل چاکنایی باز هستند. در حاشیه لازم به ذکر است که لدفوگد و مدیسون (۱۹۹۶) یک معیار اندازه‌گیری مرتبط، یعنی قله جریان هوای دهانی را به‌عنوان معتبرترین روش اندازه‌گیری انواع مختلف واک انتخاب می‌کنند. اگر از بحث آیرودینامیک سایشی‌ها به یاد داشته باشید، سرعت حجمی دهانی به مقاومت چاکنایی^۵ بستگی دارد؛ آواهای بی‌واک جریان هوای

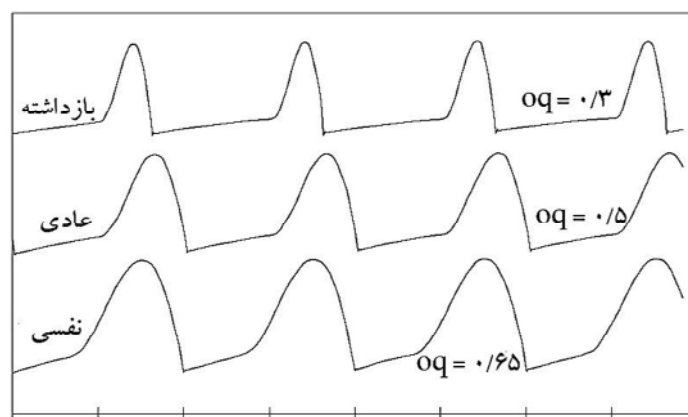
1. shutting stage
3. Bless and Abbs

2. modal voice
4. Laver

5. glottal impedence

دهانی بزرگ‌تری نسبت به آواهای واک‌دار دارند. از این رو، مقاومت چاکنایی در آواهای واک‌دار بسته به مدت‌زمانی که تارهای صوتی طی هر سیکل چاکنایی باز هستند (با فرض فشار هوای زیرحنجره‌ای مساوی)، متفاوت است.

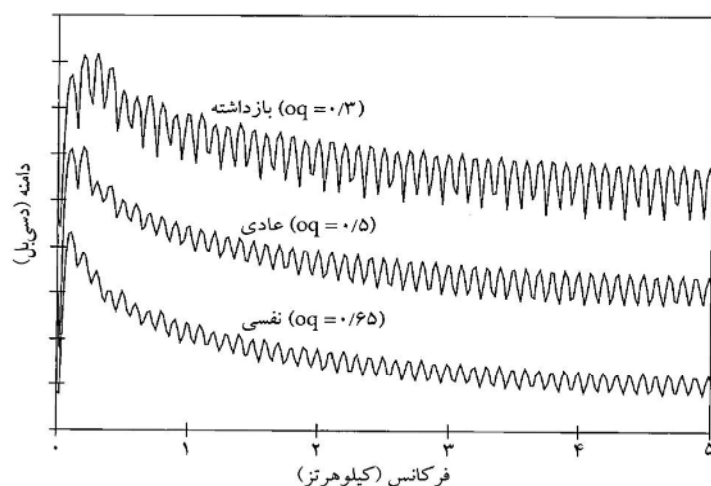
در واک عادی، تارهای صوتی در طول نیمی از یک سیکل چاکنایی بسته و در طول نیمی دیگر از آن (تقریباً) باز هستند. از این رو، نسبتی از زمان که چاکنای باز است (ضریب بازی^۱)، طی هر سیکل $0/5$ است. در واک بازداشته، تارهای صوتی، مثل دو قطعه کبد گاو، به‌گونه‌ای شُل در کنار هم قرار دارند و هوا از بین آنها بالا می‌آید. این حالت سبب طولانی‌تر شدن مرحله بسته سیکل چاکنایی و به نسبت کوتاه‌تر شدن مرحله باز آن می‌شود (در نتیجه، ضریب بازی کوچک‌تر است). در واک نفسی، تارهای صوتی ارتعاش می‌کنند، اما نه با تماس زیاد (در برخی افراد، تارهای صوتی طی واک نفسی به‌طور کامل بسته نمی‌شوند)، و بنابراین چاکنای برای زمان نسبتاً طولانی در هر سیکل باز است. شکل ۸-۲، تفاوت بین ضریب بازی بزرگ، متوسط و کوچک را در شکل‌موج‌های چاکنایی سنتز شده نشان می‌دهد. فرکانس پایه همه این شکل‌موج‌ها ۱۰۰ هرتز است.



شکل ۸-۲. شکل‌موج‌های چاکنایی که با تغییر ضریب بازی چاکنا (اُکیو) به وسیله یک سنتزکننده گفتار تولید شده‌اند. در شکل‌موج بالا، واک بازداشته با بستن چاکنا در ۷۰ درصد سیکل چاکنایی و گشودن آن در ۳۰ درصد باقی‌مانده شبیه‌سازی شده است. در شکل‌موج‌های دوم و سوم، واک عادی و نفسی به ترتیب با افزایش ضریب بازی به ۵۰ و ۶۵ درصد سیکل، شبیه‌سازی شده‌اند.

1. open quotient (oq)

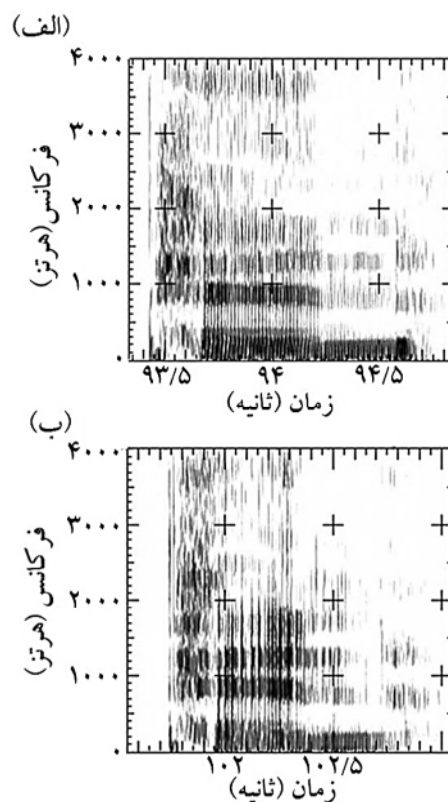
بر اساس قضیه فوریه می‌دانیم که هر تغییری در شکل موج واک (که یک موج متناوب مرکب است) منجر به یک تغییر طیفی نیز می‌شود. این در طیف واک‌های سنتز شده شکل ۸-۳ نمایان است. (قابل توجه است که اگر شکل پالس را یکسان نگه داریم و دوره تناوب را با افزایش یا کاهش مرحله بسته تغییر دهیم، فقط فاصله بین همسازها تغییر می‌کند، اما شکل طیف یکسان باقی می‌ماند.) در مقایسه با واک بازداشته یا عادی، شکل موج واک نفسی در شکل ۸-۲ شباهت بیشتری به یک موج سینوسی دارد. این، در طیف شکل ۸-۳ به صورت تسلط اولین همساز انعکاس یافته است. همه همسازهای دیگر، دامنه پایین‌تری دارند. این تسلط به این دلیل است که اولین همساز در یک موج مرکب، یک موج سینوسی است که فرکانس آن مشابه فرکانس پایه شکل موج واک است (در اینجا ۱۰۰ هرتز) و از آنجا که شکل موج واک نفسی تا حدی سینوسی است، اولین همساز تمایل به تسلط بر طیف دارد. شکل موج واک بازداشته کمترین شباهت را به یک موج سینوسی دارد و این در طیف آن انعکاس یافته است؛ به این صورت که دومین، سومین و چهارمین همساز، دامنه بالاتری نسبت به اولین همساز دارند. اختلاف بین دامنه اولین همساز با دامنه دومین همساز یک روش معتبر برای اندازه‌گیری میزان نسبی نفسی بودن یا بازداشته بودن واک به‌ویژه در واکه‌های افتاده است که در آنها سازه اول تأثیر چندانی بر دامنه هیچ‌یک از همسازها ندارد. تغییر در فرکانس پایه (F_0)، بلندی صدا و نوع واکه می‌تواند سبب تغییر دامنه نسبی اولین و دومین همساز شود. از این رو، وقتی این معیار را برای نوع واک‌سازی به کار می‌برید باید مراقب باشید. معیار طیفی دیگری که اندازه‌گیری آن در عمل دشوارتر است (به دلیل تأثیر بازخوانی دستگاه گفتار) شیب طیف است. توجه کنید که دامنه همسازها در طیف واک نفسی، نسبت به دامنه همسازها در طیف واک عادی یا بازداشته با افزایش فرکانس خیلی سریع‌تر افت می‌کند. این معیار نیز برای مقایسه انواع واک‌سازی به کار رفته است.



شکل ۸-۳. طیف قدرت شکل‌موج‌های چاکنایی که در شکل ۸-۲ نشان داده شد. برای مقایسه آسان‌تر، طیف‌ها بر روی محور عمودی کنار هم گذاشته شدند. فاصله بین علامت‌گذاری‌ها بر روی محور عمودی ۱۰ دسی‌بل است.

تفاوت‌های ذکرشده در شکل‌موج و طیف انواع واک که نتیجه تغییر در میزان ضریب بازی است، معمولاً با تفاوت‌های دیگری در سایر ویژگی‌های آکوستیکی آنها همراه است. کشش ماهیچه‌ای دخیل هنگام تولید واک بازداشته، اغلب سبب ارتعاش کندتر تارهای صوتی می‌شود و در نتیجه، واک بازداشته ممکن است فرکانس پایه کمتری در مقایسه با واک عادی یا نفسی داشته باشد. بر روی طیف‌نگاشت، پالس‌های واک طی واک بازداشته در مقایسه با واک عادی با فاصله بیشتری از یکدیگر قرار دارند (برای نمونه به شکل ۸-۴ نگاه کنید). این تفاوت، یک همبسته آکوستیکی معتبر برای واک بازداشته در زبان انگلیسی محسوب می‌شود. به همین قیاس، مدت‌زمان طولانی بازبودن چاکنای در واک نفسی سبب می‌شود در زمان کوتاهی در هر سیکل، چاکنا به حدی باز شود که جریان هوا با سرعت زیاد از آن عبور کرده و در نتیجه، مقداری سایش چاکنایی (نوفه دمشی) در هر سیکل چاکنایی واک نفسی تولید شود.

انسدادی‌ها و انسایشی‌ها ممکن است در زبان‌های مختلف بسته به نوع واک‌ی که با آن تولید می‌شوند، به گونه‌ای متمایز متفاوت باشند. این تمایزهای مبتنی بر واک می‌تواند در هر یک از سه مرحله تولید انسدادی و انسایشی روی دهد.



شکل ۸-۴. مقایسه واک عادی (الف) با واک بازداشته (ب) در دو طیف‌نگاشت برای کلمه انگلیسی "pin".

۸-۱-۲. منبع صوت در انسدادی‌ها و انسایشی‌ها

متداول‌ترین منبع انرژی آکوستیکی طی مرحله گرایش برای تولید همخوان‌های انسدادی، واک است که ممکن است واک عادی باشد یا واک نفسی یا بازداشته که کمتر اتفاق می‌افتد. انسدادی‌های پیش‌چاکنایی شده^۱ با واک بازداشته در مرحله گرایش تولید می‌شوند و انسدادی‌های پیش‌دمیده^۲ طی مرحله گرایش کمی نوفه دممش (اغتشاش تولیدشده در چاکنای) دارند.

طی مرحله بست، واک تنها منبع صوتی محتمل است و در غیر این صورت، بست بی‌واک است. علت‌های این محدودیت، هم آیرودینامیکی و هم آکوستیکی است. در

1. pre-glottalized 2. pre-aspirated

بست، وجود دمش محسوس غیرممکن است، زیرا جفت‌شدگی آکوستیکی مستقیمی بین منبع نوفه و هوای بیرون از دهان صورت نمی‌گیرد و بنابراین، دیواره‌های دستگاه گفتار هر اغتشاشی را که ممکن است طی بست تولید شود، فرو می‌نشانند. ویژگی‌های آیرودینامیکی بست انسدادی نیز از تولید دمش یا سایر منبع‌های نوفه سایش جلوگیری می‌کند. از آنجا که دستگاه گفتار به‌طور کامل طی بست انسدادی مسدود است، هوایی که از ریه‌ها جریان پیدا می‌کند، جایی برای رفتن ندارد. این وضعیت به‌ویژه برای تولید دمش که به جریان هوای زیادی نیاز دارد، مسئله‌ساز است. علاوه بر آن، از تولید هر نوفه سایشی دیگری نیز جلوگیری می‌کند. طی بست انسدادی، مقاومت آیرودینامیکی در برابر واک نیز ایجاد می‌شود، هرچند که مانع آن نمی‌شود. هوا از بین تارهای صوتی جریان پیدا می‌کند و وارد دستگاه گفتار مسدودشده می‌شود، اما خیلی زود دستگاه گفتار پر می‌شود. در نتیجه، هوای خارج‌شده از شش‌ها جایی برای رفتن ندارد و تولید واک، دیگر نمی‌تواند ادامه پیدا کند. به‌همین دلیل، در طیف‌نگاشت‌ها اغلب مقداری واک در ابتدای بست انسدادی‌ها می‌بینیم که در انتهای آن محو می‌شود. واک یا فقط در بخشی از بست باقی می‌ماند و یا گوینده برای گسترش حفره دهانی شگردی اتخاذ می‌کند و در نتیجه، مدت‌زمان تولید واک را افزایش می‌دهد. همخوان‌های مکیده^۱ (آواهای درون‌سوی^۲ چاکنایی) با یک شگرد افزایش حفره دهانی تولید می‌شوند: پایین‌آوردن حنجره (لیندو ۱۹۸۴). سایر شگردهای موجود برای گسترش حفره دهانی عبارت‌اند از ۱) جلوکشیدن ریشه زبان^۳ و ۲) کاهش کشش ماهیچه حلق یا گونه (که گسترش غیرارادی در بست را میسر می‌سازد). گاهی اوقات منشأ انسدادی‌های پیش‌خیشومی شده^۴ مثل [mb] یا [m^b] به‌لحاظ تاریخی انسدادی‌های واک‌داری است که تولید واک در بست آنها با بازشدن بینی تسهیل یافته است (توجه کنید که بستار انسداد دهانی در هر دو همخوان خیشومی و انسدادی در [mb] باقی می‌ماند). از این رو، این تولید متفاوت در مرحله بست، ممکن است ناشی از محدودیت‌های آیرودینامیکی تولید واک در طول بست انسدادی نیز باشد.

1. implosive

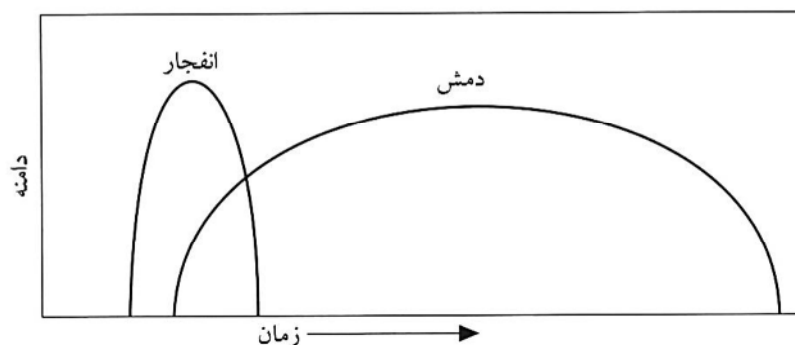
2. Ingressive

3. tongue-root advancement

4. pre-nasalized

در مرحله رهش انسدادی‌ها، انواع مختلفی از منبع صوت یافت می‌شود. اولین آنها اینکه انفجار رهش انسدادی اختیاری نیست و ویژه رهش‌های انسدادی است. انفجار انسدادی هنگامی روی می‌دهد که فشار افزایش یافته هوا در دستگاه گفتار، رهش پیدا می‌کند. هوا با سرعت زیاد از دهان خارج می‌شود و موج فشاری تولید می‌کند که فقط ۲ یا ۳ میلی‌ثانیه طول می‌کشد. بنابراین، نوبه انفجار لحظه رهش همخوان انسدادی را نشان می‌دهد. مثل مرحله گرایش، رهش نیز می‌تواند واک‌دار (با واک عادی، بازداشته یا نفسی) یا دمیده باشد.

تشخیص دو منبع صوتی در انسدادی‌های دمیده مهم است. نوبه انفجار در محل تولید همخوان تولید می‌شود (مثل نوبه سایشی) اما محل تولید نوبه دمش در چاکنای است. از این رو، صافی دستگاه گفتار برای این دو نوع منبع صوتی متفاوت است. علاوه بر آن، انفجار، دیرش بسیار کوتاهی دارد، در حالی که نوبه دمش ممکن است کاملاً بلند باشد. گرفتگی یک انسدادی برای چند میلی‌ثانیه پس از رهش بست، به قدری تنگ است که امکان ایجاد جریان هوای لازم برای تولید دمش در چاکنای نیست. در نتیجه، بلافاصله پس از رهش، شرایط برای انفجار مناسب است و نه دمش (به دلیل افزایش فشار بالای هوا و منفذ خیلی تنگ)، اما بعداً همان‌طور که گرفتگی بیشتر باز می‌شود، شرایط برای دمش مناسب می‌گردد (وقتی گرفتگی به اندازه کافی برای عبور حجم بالایی از جریان هوا باز می‌شود). در شکل ۸-۵ تقدم نسبی این دو منبع صوتی در طول مرحله رهش انسدادی‌های دمیده نشان داده شده است.



شکل ۸-۵. بازنمایی شکل کلی نوبه‌های انفجار و دمش هنگام رهش یک انسدادی دمیده.

صداهاى تولیدشده با دهان بسته

این سؤال مطرح می‌شود که چگونه می‌توان واکی را که طی بست انسدادی تولید می‌شود، بشنویم و بر روی طیف‌نگاشت ببینیم. واکی که طی بست انسدادی می‌شنویم، مانند موسیقی است که وقتی همسایه طبقه بالا صدای استریوی خود را بلند می‌کند، می‌شنویم. دقیقاً مثل واک انسدادی، ممکن است مجرای هوای مستقیمی برای انتقال صدا نباشد (جزء اینکه پنجره‌های هر دو طبقه بالا و پایین باز باشد)، اما شما صدای استریو را می‌شنوید، زیرا کف اتاق (سقف شما) موسیقی را در فرکانس‌های بازخوانی کف اتاق انتقال می‌دهد. اما به هر حال، این موسیقی که دیوار را به لرزش درآورده، همان صدایی نیست که اگر داخل اتاق و کنار بلندگوها بودید، می‌شنیدید. همه آن چیزی که می‌شنوید، بوم، بوم، بوم نت^۱ بم^۱ است. ضربه‌ها را می‌شنوید، اما ملودی را نه. زیرا فرکانس‌های بازخوانی سقف‌ها و دیوارها پایین است (زیرا آنها اجسام حجیم و بزرگ هستند) و بنابراین، در نت‌های پایین به‌خوبی تشدید یا بازخوان می‌شوند. به همین قیاس، هنگامی که تارهای صوتی در طول یک بست انسدادی ارتعاش می‌کنند، دیواره‌های دستگاه گفتار، صوت را منتقل می‌کنند، اما فقط مؤلفه‌های با فرکانس پایین را. از این رو، در طیف‌نگاشت‌ها، واک در طول بست انسدادی به‌صورت یک نوار با فرکانس پایین به نام خط واک^۲ در پایین طیف‌نگاشت ظاهر می‌شود.

شیوه‌های بسیاری برای ایجاد انفجار رهش انسدادی وجود دارد. شناخته‌شده‌ترین نوع انفجار رهش در انسدادی‌های ششی دیده می‌شود، همان‌طور که در زبان انگلیسی داریم. در همخوان‌های فورانی^۳ (آواهای برون‌سوی^۴ چاکنایی) فشار هوای پشت بست که با بستن چاکنای و بالا آوردن حنجره در طول بست ایجاد

1. bass

2. voice bar

3. ejective

4. egressive

می‌شود، معمولاً نسبت به فشار هوا در انسدادی‌های ششی بیشتر است و انفجار حاصل دامنه انرژی بیشتری دارد. این، به‌ویژه در مورد فورانی‌های بدنه‌ای^۱ صادق است، زیرا فشار بالای هوای دهانی تنها می‌تواند در فورانی‌های بدنه‌ای تولید شود تا در فورانی‌هایی که محل‌های تولید دیگری دارند، زیرا حفره دهان در [k] نسبت به [p] کوچک‌تر است. از این رو، شاید تعجب‌آور نباشد که فورانی‌های بدنه‌ای در زبان‌های دنیا نسبت به سایر فورانی‌ها متداول‌ترند. (جان او‌هالا به این استدلال درباره فراگیر بودن همخوان فورانی [kʰ] در زبان‌های دنیا ایراد وارد می‌کند و معتقد است که انفجار رهش انسدادی‌های پسین معمولاً به‌لحاظ درکی، برجسته‌تر از انسدادی‌های پیشین است، زیرا دستگاه گفتار در جلوی گرفتگی بلندتر است و بنابراین انفجار، حاوی اطلاعات بیشتری درباره تولید است.) علاوه بر آن، همخوان‌های فورانی با دو انفجار رهش توصیف می‌شوند: رهش دهانی و رهش چاکنایی (لیندو ۱۹۸۴، مک‌دانو و لدفوکد ۱۹۹۳). از این رو، گاهی اوقات برای آوانویسی این انسدادی‌های «چاکنایی‌شده»^۲ از یک انسدادی چاکنایی پس از علامت انسدادی دهانی استفاده می‌شود، مثل [kʰ] که در آی‌پی‌ای به‌صورت [kʰ] است.

از سوی دیگر، انفجار رهش همخوان‌های مکیده (آواهای درون‌سوی چاکنایی) ضعیف‌تر از انسدادی‌های ششی است. این، احتمالاً بدان جهت است که تراکم فشار هوای دهانی در بست همخوان‌های مکیده، خیلی بزرگ نیست. توجه به این نکته مهم است که انفجار ضعیف، لزوماً نشان‌دهنده جریان هوای درون‌سو نیست، بلکه صرفاً نشان‌دهنده این است که همخوان‌های مکیده تمایل چندانی برای ایجاد تغییر در فشار هوای دهانی طی بست انسدادی ندارند.

انفجار رهش **نچ‌آواها** نوعاً بسیار بلندتر از هر نوع انفجار رهش دیگری است، زیرا اختلاف فشار هوای دهان با فشار اتمسفری در نچ‌آواها نسبتاً بزرگ است. از آنجا که نچ‌آواها به‌وسیله یک حفره خیلی کوچک هوا بین زبان و سقف دهان تولید می‌شوند،

1. dorsal ejective

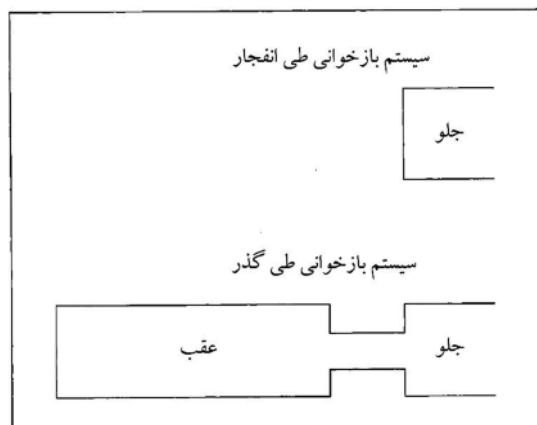
2. glottalized stops

یک جابه‌جایی نسبتاً کوچک زبان سبب یک تغییر خیلی بزرگ در فشار هوای دهانی می‌شود و در نتیجه، امکان تولید یک انفجار رهش درون‌سوی بلند را میسر می‌سازد.

۸-۲. توابع صافی دستگاه گفتار در انسدادی‌ها

دو نوع صافی در دستگاه گفتار بر ویژگی‌های آکوستیکی همخوان‌های انسدادی تأثیرگذار است. نخست، شکل‌بندی دستگاه گفتار بلافاصله پس از بست که ویژگی‌های طیفی انفجار رهش را شکل می‌دهد. دوم، شکل‌بندی و حرکت‌های دستگاه گفتار قبل و بعد از بست انسدادی که شکل‌دهنده گذرهای سازه‌ای در دو مرحله گرایش و رهش تولید انسدادی است.

انفجار رهش در اصل یک موج گذراست که یک طیف مسطح دارد (رجوع کنید به فصل ۱). این منبع نوفه به‌وسیله بازخوان‌های بخشی از دستگاه گفتار که در جلوی بست انسدادی است، شکل می‌گیرد، زیرا بست در زمان رهش هنوز کاملاً تنگ است و در نتیجه، حفره‌های جلو و عقب دستگاه گفتار به‌لحاظ آکوستیکی جفت نشده‌اند (شکل ۸-۶). از این رو، انفجار رهش به‌لحاظ آکوستیکی شبیه به همخوان‌های سایشی است. انسدادی‌های لبی، حفره جلویی ندارند، بنابراین طیف آنها به‌وسیله آکوستیک منبع صوت بدون هیچ‌گونه فیلتری در دستگاه گفتار شکل می‌گیرد. بدین ترتیب، انفجار رهش انسدادی‌های لبی هیچ قله سازه‌ای ندارد و انرژی به‌صورت پراکنده بر روی همه فرکانس‌ها پخش است. انسدادی‌های دندانی و لثوی یک حفره جلویی کوتاه دارند و بنابراین، قله‌هایی با فرکانس بالا در انرژی طیفی دارند. انسدادی‌های سخت‌کامی و نرم‌کامی، حفره جلویی بلندتری دارند و بنابراین، طیف آنها قله‌هایی با فرکانس پایین‌تر دارد و انفجار آنها معمولاً ساختار سازه‌ای بیشتری نسبت به سایر انفجارهای انسدادی دارد.

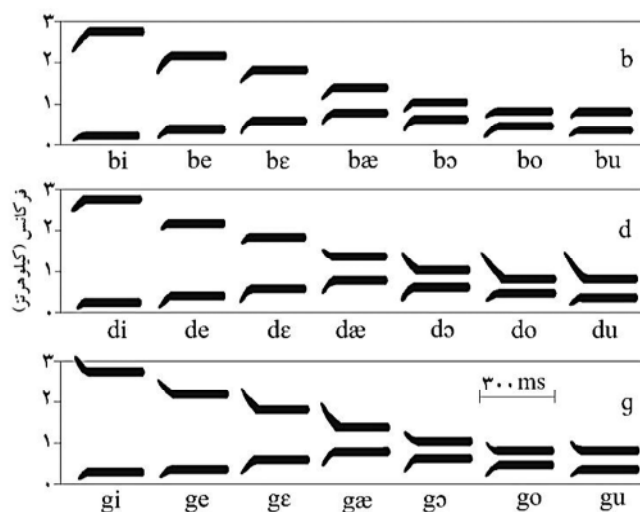


شکل ۸-۶. مدل لوله آکوستیکی انفجار و رهش در انسدادی‌ها.

در مرحله گرایش و نیز در مرحله رهش انسدادی‌ها پس از انفجار، حفره‌های جلو و عقب دستگاه گفتار به لحاظ آکوستیکی با یکدیگر جفت می‌شوند و منبع صوت هم برای واک و هم برای دمش در چاکنای است. از این رو، ویژگی‌های آکوستیکی این دو مرحله در تولید همخوان‌های انسدادی با مدل‌هایی که برای بحث درباره مقادیر سازه‌ای واکه‌ها به کار بردیم (نظریه اختلال و مدل‌های دولوله‌ای)، بهتر فهمیده می‌شود. گرایش و رهش، همان‌طور که از اسم آنها برمی‌آید، حرکت هستند نه بستر. از این رو، مثل واکه‌های مرکب، اطلاعاتی که درباره محل تولید دارند، در گذرهای سازه‌ای نهفته است و نه در مقادیر معین سازه‌ای و البته خصوصیات یک حرکت به نقطه آغاز و پایان آن نیز بستگی دارد. این در شکل ۸-۷ نشان داده شده است.

دیلتز و همکاران (۱۹۵۵) این الگوهای سازه‌ای را برای کنترل برون‌داد یک سترکننده گفتار به کار بردند. الگوها، وضعیت ایستا و گذرهای سازه‌ای مورد نیاز برای تولید هجاهای فهرست‌شده در شکل ۸-۷ را نشان می‌دهند. در فصل ۶ گفتیم که سازه دوم (بازخوان دستگاه گفتار) در واکه‌های پیشین بالاست و اولین سازه در واکه‌های افتاده بالاست. نکته قابل توجه در این شکل این است که حرکت‌های سازه‌ای (یا گذرها) که در درک محل تولید یک انسدادی مشارکت دارند، به نوع واکه بعدی بستگی دارند. این، به‌ویژه در مورد [d] بسیار چشمگیر است: وقتی F_2 واکه

بالاست، گذر F_2 نیز افزایش پیدا می‌کند، اما وقتی F_2 واکه پایین است، گذر سازه نیز افت می‌کند. «دیلتر» و همکاران دریافتند که بخشی از گذر که به‌لحاظ درکی مهم است، در واقع در سیگنال آکوستیکی حضور ندارد، اما می‌توان آن را از مجموعه‌ای از هجاها مثل هجاهای شکل ۸-۷ به‌دست آورد. برای مثال، اگر منحنی‌های گذر F_2^1 در هجاها را روی هم بگذارید، مثلاً با [d] شروع کنید، سپس منحنی‌های گذر F_2 را در زمان به عقب امتداد دهید، به نظر می‌رسد که در حدود ۱۸۰۰ هرتز با هم تلاقی می‌کنند. دیلتر و همکاران، این فرکانس آغازی فرضی را **فرکانس کانونی**^۲ F_2^2 نامیدند. در مطالعه آنها، فرکانس کانونی F_2 برای همخوان‌های لبی در ۷۲۰ هرتز بود، اما همخوان‌های نرم‌کامی، کانون F_2 بالاتری، در ۳۰۰۰ هرتز داشتند. نتایج آنها در مورد نرم‌کامی‌ها نادرست است، زیرا آنها از نمونه‌های سنتز شده‌ای استفاده می‌کردند که فقط دو سازه داشتند، اما الگو درست است. فرکانس کانونی همخوان‌های لبی پایین، فرکانس کانونی لثوی‌ها، متوسط و نرم‌کامی‌ها فرکانس کانونی نسبتاً بالاتری دارند.



شکل ۸-۷. الگوهای گذر F_1 و F_2 در رهش که برای سنتز [b]، [d] و [g] در بافت قبل از واکه‌ها به کار رفت. اقتباس از دیلتر و همکاران (۱۹۵۵، ص. ۷۷۰) و انتشار با کسب اجازه.

1. trajectory 2. locus frequency

این واقعیت که کانون در عمل امتداد ندارد، ناشی از تغییراتی است که با جفت‌شدگی آکوستیکی حفره‌های جلو و عقب طی تولید انسدادی روی می‌دهد. همان‌طور که قبلاً گفته شد، وقتی حفره‌های جلو و عقب جفت می‌شوند، ویژگی‌های آکوستیکی دستگاه گفتار به‌طور ناگهانی تغییر می‌کند. در لحظه رهش انسدادی، حفره‌های جلو و عقب جفت نیستند، در نتیجه فرکانس‌های بازخوانی حفره عقب تأثیری بر طیف گفتار ندارند، اما همین‌که بست بیشتر باز می‌شود، حفره‌های جلو و عقب جفت می‌شوند. به‌همین دلیل انفجار رهش [d] بازخوان‌هایی با فرکانس بالاتر نسبت به انفجار رهش [g] دارد، در حالی که [g] کانون F_2 بالاتری نسبت به [d] دارد. انفجار رهش [d] به‌وسیله بازخوان حفره جلویی (که فرکانس بالایی دارد) شکل می‌گیرد، اما کانون F_2 مربوط به بازخوان حفره عقبی است. هم انفجار و هم کانون F_2 در [g] مربوط به بازخوانی در حفره جلویی است.

تغییر سریع در جفت‌شدگی حفره‌های عقب و جلو از سویی توجیه می‌کند که چرا همه همخوان‌های انسدادی، یک F_1 با فرکانس پایین نزدیک ناحیه بست دارند. طی بست، دو حفره جفت نیستند، در نتیجه سازه اول وجود ندارد (زیرا سازه اول یک بازخوان «هلمهولتز» است، به فصل ۶ رجوع کنید). بازخوان «هلمهولتز» که F_1 را تولید می‌کند با یک گردن بسیار تنگ (و در نتیجه، با یک F_1 پایین) فعال می‌شود. همان‌طور که گرفتگی بازتر می‌شود، گردن بازخوان «هلمهولتز» نیز بازتر می‌شود و در نتیجه، فرکانس F_1 افزایش پیدا می‌کند. علاوه بر آن، همان‌طور که بست رهش می‌کند و گرفتگی بازتر می‌شود، حفره‌های جلو و عقب به‌لحاظ آکوستیکی جفت می‌شوند. اما طی مرحله گرایش، عکس این روی می‌دهد: بدین‌صورت که اولین سازه با فرکانس نسبتاً بالایی آغاز می‌شود و همان‌طور که گرفتگی شکل می‌گیرد، افت می‌کند.

گذرهای سازه‌ای

در همخوان‌های لبی (دولبی) فرکانس همه سازه‌ها نزدیک بست انسدادی پایین است. بر اساس نظریه اختلال، دلیل آن برای همه بازخوان‌های دستگاه گفتار، ایجاد گرفتگی در محل یک گره سرعت (لب‌ها) است. کانون F_2 در انسدادی‌های تیغه‌ای (دندانی و لثوی) اندکی بالاتر از دومین فرکانس بازخوانی یک لوله واحد است. فرکانس F_3 نزدیک بست انسدادی‌های تیغه‌ای نیز نسبت به یک لوله واحد بالاتر است. این تأثیرات با یک مدل دولوله‌ای که یک گرفتگی تقریباً در جلوی دستگاه گفتار دارد، قابل پیش‌بینی است (شکل ۶-۴). در انسدادی‌های بدنه‌ای (سخت‌کامی و نرم‌کامی)، نزدیک بست انسدادی، هم‌گرایی F_2 و F_3 دیده می‌شود، زیرا محل بست، نزدیک محل تلاقی بازخوان‌های حفره جلو و عقب است. البته کانون سازه‌ای واجگونه‌های همخوان‌های نرم‌کامی انگلیسی آمریکایی (در جایگاه قبل از واکه‌های پیشین و پسین) متفاوت است.

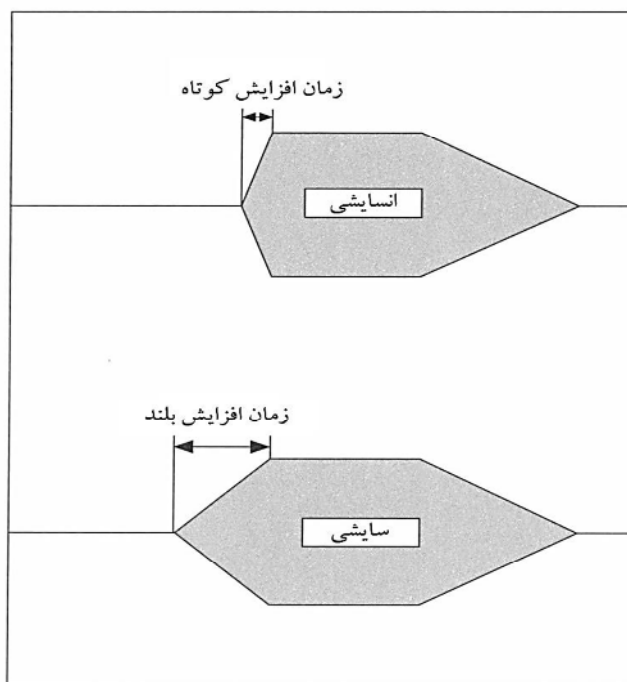
۸-۳. انسایشی‌ها

در مرحله رهش یک انسدادی ممکن است به جای واک یا دمش، سایش تولید شود. این نوع انسدادی «سایشی‌شده»، انسایشی نامیده می‌شود. نوفه سایش انسایشی‌ها معمولاً در همان محلی تولید می‌شود که انسداد آنها تولید می‌شود (با مقداری فضا برای ایجاد تمایز، همان‌طور که انسایشی‌های [ts] و [ʃ] نشان می‌دهند). اما این شیوه تولید در هر حال، ویژگی توصیف‌کننده انسایشی‌ها محسوب نمی‌شود، زیرا انسایشی‌های ناهم‌اندام^۱ نیز در زبان‌ها وجود دارند. به‌عنوان مثال، مک‌دانو و لدفورد (۱۹۹۳) دریافتند که انسدادی‌هایی که به صورت [tʰ] در زبان ناواهو آوانگاری می‌شوند، به لحاظ آکوستیکی و واجی، انسایشی هستند. سایشی‌های ناهم‌اندام در زبان سوتوی شمالی^۲ و زبان‌های مرتبط با آن یافت می‌شود. مهم‌ترین تمایز آکوستیکی بین یک انسایشی و توالی از یک انسدادی و یک سایشی در این است که دامنه نوفه

1. heterorganic

۲. Northern Sotho. سوتوی شمالی زبانی در آفریقای جنوبی است. (مترجم)

سایش در انسایشی‌ها خیلی سریع به دامنه کامل می‌رسد، اما در سایشی‌ها کندتر. این ویژگی آوایی که در شکل ۸-۸ نشان داده شده، زمان افزایش^۱ نامیده شده است.



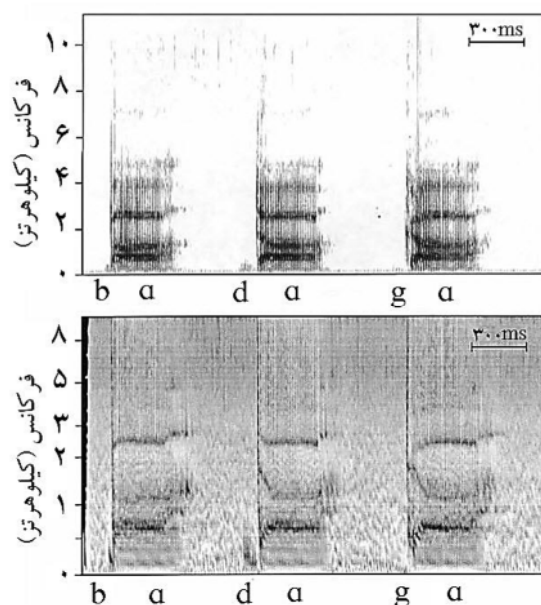
شکل ۸-۸. بازنمایی تفاوت بین پوش شکل موج انسایشی با سایشی. در انسایشی‌ها دامنه به سرعت به اوج می‌رسد، اما در سایشی‌ها کندتر.

۸-۴. ویژگی‌های شنیداری انسدادی‌ها

شکل ۸-۹، طیف‌نگاشت و حلزون‌نگاشت سه هجای بی‌معنای [ba]، [da] و [ga] را نشان می‌دهد که توسط یک گویشور انگلیسی‌زبان تولید شده‌اند. در این شکل، تعدادی از تفاوت‌های مهم بین بازنمایی‌های آکوستیکی با شنیداری انسدادی‌ها قابل مشاهده است. نخست، همان‌طور که در فصل ۶ دیدیم، ناحیه F_2 در بازنمایی شنیداری تا حدی نسبت به طیف‌نگاشت برجسته‌تر است (با این وجود، هر دو شکل فرکانس‌های از صفر تا ۱۱ کیلوهرتز را نمایش می‌دهند). این در حلزون‌نگاشت‌های شکل ۸-۹ در حرکات بزرگ F_2

1. rise time

پس از رهش انسدادی‌ها واضح است. از این رو، هرچند که ممکن است گذرهای F_2 تا حدی در طیف‌نگاشت‌ها قابل تمایز نباشند (به‌ویژه وقتی یک محدوده فرکانسی بزرگ نشان داده می‌شود) اما، در بازنمایی شنیداری تقویت می‌شوند. از سویی محاسبه کانون F_2 که در بالا بحث کردیم، بر اساس منحنی‌های گذر آکوستیکی سازه‌ها صورت گرفت. احتمالاً از آنجا که مقیاس فرکانس شنیداری غیرخطی است، فرکانس‌های کانونی شنیداری F_2 نیز متفاوت از مقادیر آکوستیکی آن خواهد بود. من هیچ تحقیقی که به‌طور مفصل به این موضوع پرداخته باشد، نمی‌شناسم، اما این نکته‌ای است که ممکن است بر درک ما از مفهوم «فرکانس کانونی» تأثیر داشته باشد.



شکل ۸-۹. طیف‌نگاشت‌ها (بالا) و حلزون‌نگاشت‌های (پایین) هجاهای [ba]، [da] و [ga] که توسط یک گویشور مرد انگلیسی‌زبان تولید شده‌اند، برخی تفاوت‌های بین بازنمایی آکوستیکی با شنیداری در همخوان‌های انسدادی را نشان می‌دهند.

دوم اینکه، انفجارهای رهش نیز در حلزون‌نگاشت‌ها در مقایسه با طیف‌نگاشت‌ها برجسته‌تر هستند. این نتیجه مکانیسم کنترل خودکار بهره^۱ است که

1. automatic gain control mechanism

به وسیله استتار سیستم شنوایی محیطی ایجاد می شود (رجوع کنید به بخش ۴-۴). پاسخ سیستم شنوایی پس از یک سکوت کوتاه، قوی تر است نسبت به وقتی که صدا پیوسته است. از این رو، آغازه ها به طور کلی پاسخ بزرگی در سیستم شنوایی ایجاد می کنند. در انسدادی ها، این عملکرد سیستم شنوایی بر اهمیت انفجار رهش می افزاید. اما توجه داشته باشید که حتی هجاهایی که با واکه آغاز می شوند، نیز ممکن است «انفجار رهش» در سیستم شنوایی ایجاد کنند. این واقعیت می تواند توجیهی باشد برای اینکه چرا هجاهای سنتز شده «بی انفجار» به نظر می رسد با یک همخوان انسدادی آغاز می شوند؛ سیستم شنوایی یک انفجار تولید می کند، هرچند که سیگنال آکوستیکی فاقد آن است.

۵-۸. درک انسدادی ها در بافت واکه ای

همخوان انسدادی نرم کامی [k] وقتی پس از واکه [i] تولید شود، اغلب به انسایشی لثوی کامی [ʃ] تبدیل می شود. این تغییر آوایی که به صورت زیر نوشته می شود، کامی شدگی^۱ نامیده می شود و در زبان های بسیاری روی می دهد:

$$K \longrightarrow \text{ʃ} / _ i$$

گیون^۲ (۱۹۹۸) به مواردی از کامی شدگی در اسلاوی، هندوایرانی، سالیش^۳، بانتو^۴، انگلیسی، مایان^۵ و چینی اشاره می کند.

با بررسی دقیق تر به نظر می رسد که کامی شدگی علت تولیدی دارد. همخوان [k] پیشین می شود، زیرا آوای بعدی، واکه پیشین [i] است. از آنجا که در زبان های بسیاری، بافت سایر واکه های پیشین مثل [e] و [æ] نیز شرط وقوع این تغییر آوایی

1. palatalization

2. Guion

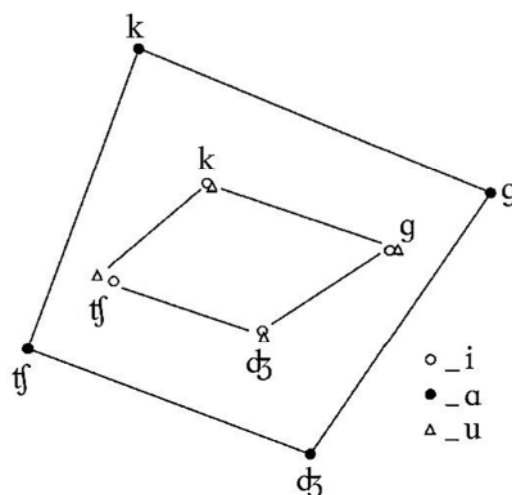
۳. Salish. سالیش، یکی از خانواده های زبانی بومیان آمریکا و کانادا در سواحل اقیانوس آرام است. (مترجم)

۴. Bantu. یکی از زبان های خانواده زبانی نیجر-کنگو است که در جنوب صحرای آفریقا به آن صحبت می شود. (مترجم)

۵. Mayan. زبان مایاها که در آمریکای جنوبی می زیستند. (مترجم)

است، این تبیین قوی‌تر می‌شود. اما به هر حال، کمی عجیب به نظر می‌رسد که حاصل کامی‌شدگی، یعنی [ʃ] محل تولیدی جلوتر نسبت به [i] که بافت کامی‌شدگی است، دارد. اگر کامی‌شدگی یک فرایند کاملاً تولیدی است که در نتیجه آن [k] شباهت بیشتری به [i] پیدا می‌کند، پس انتظار می‌رود که حاصل آن، همخوان کامی [c] باشد. چرا این گونه نیست؟

آنچه در اینجا اتفاق افتاده به این دلیل است که درک نیز در کامی‌شدگی نقشی ایفا می‌کند. یکی از بهترین راه‌های اثبات آن با استفاده از ماتریس ابهام گیون (۱۹۹۸) است. شکل ۸-۱۰ نتیجه چند تحلیل مقیاس‌گذاری چندبعدی (ام‌دی‌اس) برای محاسبه ابهام بین [k]، [g]، [ʃ] و [dʒ] را نشان می‌دهد. این آواها توسط گروهی از گویشوران انگلیسی آمریکایی تولید و توسط گروهی دیگر از شنوندگان انگلیسی آمریکایی درک شده‌اند. هجاهای CV مورد آزمایش با مقداری نوفه پخش شدند (با نسبت سیگنال به نوفه ۲+ دسی‌بل) و شنوندگان همخوان آغازی هجاها را به صورت “k”، “g”، “ch” یا “j” شناسایی کردند.



شکل ۸-۱۰. نقشه‌های درکی [k]، [g]، [ʃ] و [dʒ] در بافت‌های واکه‌ای مختلف. این نقشه‌ها با استفاده از ماتریس‌های ابهام گیون (۱۹۹۸) رسم شده‌اند.

در شکل ۸-۱۰ فضای درکی به نحوی شکل گرفته که انسدادی‌های کامی در بالا و انسایشی‌های لثوی کامی در پایین آن قرار دارند. همخوان‌های بی‌واک در سمت چپ بُعد افقی و همخوان‌های واک‌دار در سمت راست آن قرار دارند. شناسایی این همخوان‌ها برای شنوندگان خیلی آسان‌تر است وقتی که واکه بعدی [a] باشد. به همین دلیل مربع حاصل از اتصال نقطه‌های توپر (بافت a_) خیلی بزرگ‌تر است نسبت به فضایی که با اتصال نقطه‌های دو بافت i_ و u_ به دست آمده است. وقتی واکه پس از همخوان انسدادی، [i] یا [u] باشد، فضای درکی کاهش پیدا می‌کند.

این نتیجه، چه چیزی را درباره کامی‌شدگی به ما می‌گوید؟ من فکر می‌کنم کاهش فاصله درکی بین [k] و [ʔ] در بافت [i_] به این معناست که وقتی یک گویشور [ʔ] را به جای [k] در این بافت تولید می‌کند، احتمال کمی دارد که شنوندگان متوجه آن بشوند. علت تولیدی همچنان پابرجاست: [i] فشار پیشین‌شدگی اعمال می‌کند. اما یک اثر «جواز» درکی^۱ نیز وجود دارد، بدین صورت که تغییر [k] به [ʔ] در بافت [i] مجاز است، زیرا این تغییر چندان قابل توجه به نظر نمی‌رسد.

1. perceptual licensing

منابعی برای مطالعه بیشتر

- Bless, D. M. and Abbs, J. H. (1983) *Vocal Fold Physiology: Contemporary research and clinical issues*, San Diego: College Hill Press.
- مجموعه مقاله‌هایی در زمینه واک و واک‌سازی.
- Delattre, P. C., Liberman, A. M., and Cooper, F. S. (1955) Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 27, 769-73.
- مقاله کلاسیکی که ایده «کانون سازه‌ای» را به‌عنوان یک پارامتر آکوستیکی مهم برای محل تولید همخوان‌ها مطرح کرد.
- Guion, S. G. (1998) The role of perception in the sound change of velar palatalization. *Phonetica*, 55, 18-52.
- مطالعه‌ای مهم در مورد درک محل تولید همخوان‌های انسدادی.
- Ladefoged, P. and Maddieson, I. (1996) *The Sounds of the World's Languages*, Oxford: Blackwell.
- مجموعه‌ای ارزشمند در زمینه آواشناسی، حاوی مطالبی خواندنی و معتبر در زمینه مقیاس‌های مختلف انواع واک‌سازی.
- Layer, J. (1980) *The Phonetic Description of Voice Quality*, Cambridge: Cambridge University Press.
- طبقه‌بندی کامل و خردمندانه انواع واک‌سازی.
- Lindau, M. (1984) Phonetic differences in glottalic consonants. *Journal of Phonetics*, 12, 147-55.
- مقایسه ویژگی‌های آکوستیکی انسدادی‌های فورانی در زبان‌های دنیا.
- McDonough, J. and Ladefoged, p. (1993) Navajo stops. *UCLA Working Papers in Phonetics*, 84, 151-64.
- مطالعه آکوستیکی انسدادی‌های ناواهو، به‌همراه گزارش تولید وی‌آئی با زمان بسیار بلند و وی‌آئی فورانی در این زبان.

تمرین

واژگان تخصصی

این اصطلاحات را تعریف کنید: سه مرحله تولید انسدادی، انواع واک‌سازی، واک عادی، ضریب بازی، واک بازداشته، واک نفسی، نوفه دمش، انسدادی‌های پیش‌چاکنایی شده، انسدادی‌های پیش‌دمیده، مکیده‌ها، انسدادی‌های پیش‌خیشومی شده، خط واک، انفجار رهش انسدادی، فورانی‌ها، دهانی، انسدادی‌های چاکنایی شده، نیچ‌آواها، گذرهای سازه‌ای، فرکانس کانونی، زمان افزایش.

پرسش

۱. کجی طیفی هر یک از طیف‌های شکل ۸-۳ را اندازه‌گیری کنید. در هر طیف، مماس بر قله‌های همسازها یک خط صاف با دقت کامل رسم کنید. سپس کجی را به صورت تغییرات دامنه در ۱۰۰۰ هرتز اندازه بگیرید.
۲. با استفاده از طیف‌نگاشت‌های شکل ۸-۷، فرکانس کانونی F_2 را برای همخوان [d] به دست آورید. برای این کار، شکل را از راست به چپ مثل یک آکاردئون یا یک آنتن رادیویی روی هم تا کنید به نحوی که همه هجاها روی [di] قرار بگیرند. من از خط‌کش برای علامت‌گذاری دو نقطه در هر هجا استفاده کردم: نقطه F_2 در شروع گذر و نقطه F_2 در پایان گذر، با تقریباً همان فاصله افقی که در ابتدا از یکدیگر داشتند. با اتصال این دو نقطه، گذرهای سازه‌ای برای هر هجای CV بازسازی می‌شود. برای به دست آوردن فرکانس کانونی، خطوط را در زمان به عقب امتداد دهید. آنها باید با یکدیگر تلاقی کنند (نزدیک علامت‌گذاری‌های محور فرکانس) و نقطه تلاقی، کانون F_2 برای این نمونه‌های [d] است.
۳. اگر در تعداد نامتناهی از زبان‌ها، همخوان [k] وجود داشته باشد، اما [p] وجود نداشته باشد (مثلاً عربی)، این چگونه به دیدگاه او‌هالا در مورد نادربودن فورانی‌های نرم‌کامی ارتباط پیدا می‌کند.

۴. آیا انسایشی‌ها انفجار رهش انسدادی دارند؟ چگونه ممکن است زمان افزایش کوتاه منجر به کاهش برجستگی شنیداری انفجار رهش انسایشی‌ها شود؟
۵. در زبان‌های بسیاری، تقابل‌های انسدادی بیشتر در آغاز هجا (قبل از واکه) مجاز است تا در پایانه هجا. به این، «جواز آغاز»^۱ می‌گویند. به نظر شما کدام یک از تقابل‌های زیر در جایگاه پایانه محدودتر است: تقابل‌های محل تولید (دولبی در مقابل دندانی، و غیره) یا تقابل‌های چاکنایی (دمیده در مقابل نادمیده)؟ در این رابطه بررسی‌ها نشان می‌دهد که انسدادی‌های آغاز همیشه رهش می‌کنند در حالی که همخوان‌های پایانه ممکن است رهش کنند یا نکنند.

فصل ۹

خیشومی‌ها و کناری‌ها

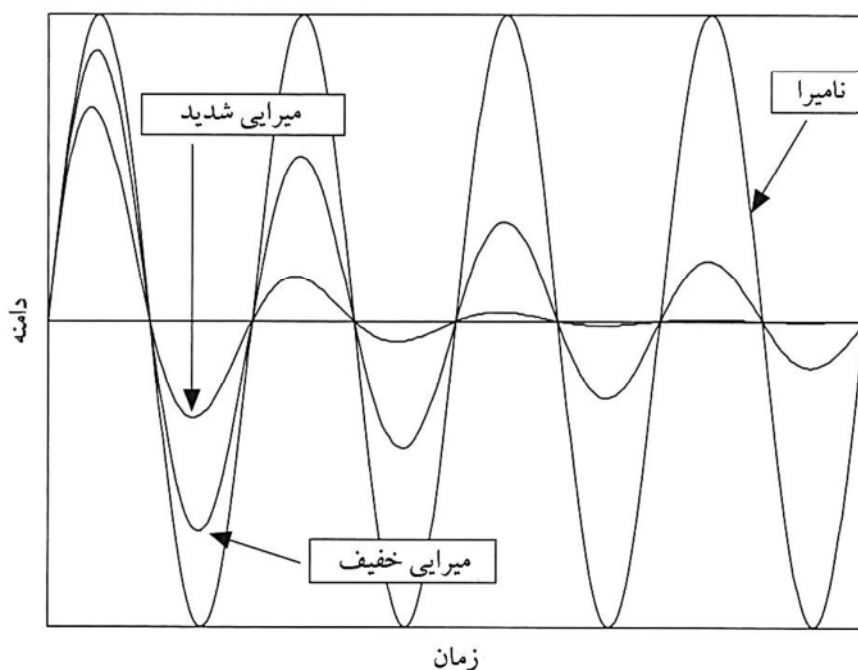
در فصل ۸ دیدیم که انسدادی‌ها و انسایشی‌ها بسیار پیچیده‌تر از واکه‌ها و سایشی‌ها هستند، زیرا در تولید آنها بیش از یک مرحله و بیش از یک نوع منبع صوتی وجود دارد. خیشومی‌ها و کناری‌ها نیز پیچیده‌تر از واکه‌ها و سایشی‌ها هستند، اما به شکلی دیگر: ویژگی‌های صافی دستگاه گفتار در آنها بسیار پیچیده‌تر است. مانند فصل ۸، با مختصری مباحث فرعی (اما مرتبط با موضوع) شروع می‌کنیم.

۹-۱. پهنای باند

یکی از تفاوت‌های تابع فیلتر دستگاه گفتار در خیشومی‌ها در مقایسه با واکه‌های دهانی این است که پهنای قلهٔ بازخوانی (پهنای باند) سازهٔ اول در خیشومی‌ها بزرگ‌تر است. از این رو، پیش از پرداختن به ویژگی‌های آکوستیکی خیشومی‌ها، به بحث دربارهٔ پهنای باند سازه‌ها می‌پردازیم.

شکل ۹-۱، یک موج سینوسی نامیرا و دو موج سینوسی میرا را نشان می‌دهد.

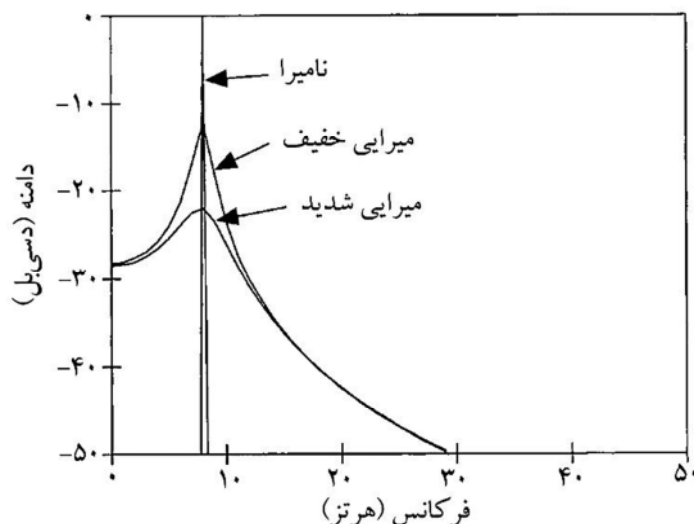
دامنهٔ امواج سینوسی میرا با گذشت زمان کاهش پیدا می‌کند. این کاهشِ دامنه شبیه هل دادن یک کودک بر روی یک تاب است: تاب را هل می‌دهید، اما خیلی طول نمی‌کشد که از حرکت باز می‌ایستد، زیرا انرژی‌ای که موقع هل دادن به تاب وارد می‌کنید با مقاومت طبیعی که در برابر حرکت تاب بر اثر کشش جاذبهٔ زمین و اصطکاک که در طی حرکت تاب و (بچه) در فضا به وجود می‌آید، از بین می‌رود. دامنهٔ موج سینوسی میرایی که با عنوان «میرایی شدید» در شکل مشخص شده، بسیار سریع‌تر از موجی که با عنوان «میرایی خفیف» مشخص شده، از بین می‌رود. موجی که میرایی آن شدید است، شبیه تاب خوردن بر روی کرهٔ زمین است و موجی که میرایی آن خفیف‌تر است، شبیه تاب خوردن بر روی کرهٔ ماه (که جاذبهٔ آن ضعیف‌تر و تراکم اتمسفر در آنجا کمتر است).



شکل ۹-۱. مقایسهٔ یک موج سینوسی نامیرا با دو موج سینوسی میرا. امواج دارای فرکانس و فاز یکسان هستند.

گفتیم که طیف موج سینوسی، دامنه و فرکانس آن را نشان می‌دهد. با توجه به این واقعیت که موج سینوسی میرا، شکلی کاملاً سینوسی ندارد، قابل پیش‌بینی است که طیف این امواج، پیچیده‌تر از امواج سینوسی نامیراست. شکل ۹-۲ نشان می‌دهد که قله طیف امواج سینوسی میرا در همان فرکانس‌هایی است که قله موج سینوسی نامیراست، اما با این تفاوت که در امواج میرا، انرژی روی سایر فرکانس‌های مجاور قله نیز پخش شده است. پیامد طیفی میرایی موج سینوسی، گسترش قله حول فرکانس موج سینوسی است. در یک موج سینوسی نامیرا، قله بی‌نهایت باریک است، اما هرچه میرایی بیشتر و بیشتر می‌شود، قله پهن‌تر و پهن‌تر می‌گردد. دوباره به شکل موج‌های شکل ۹-۱ نگاه کنید. موجی که مانند تاب‌خوردن روی ماه است (با میرایی خفیف)، در مقایسه با موج دیگر، بیشتر به یک موج سینوسی خالص شباهت دارد. از آنجا که شکل موج «میرای خفیف» شباهت بیشتری به یک موج سینوسی دارد تا شکل موج «میرای شدید»، طیف آن نیز شباهت بیشتری به طیف یک موج سینوسی دارد. به عبارت دیگر، قله آن باریک‌تر است. موجی که با سرعت بیشتری تحلیل می‌رود، هم شکل و هم طیف آن شباهت کمتری به یک موج سینوسی دارد.

از آنجا که دیواره‌های دستگاه گفتار نرم هستند، مقداری از انرژی صوتی حاصل از ارتعاش چاکنا را جذب می‌کنند (اینرسی هوای درون و بیرون دستگاه گفتار نیز مقداری از انرژی صوتی را جذب می‌کند). ارتعاش امواج فشار صوتی که در دستگاه گفتار تشدید یا بازخوان می‌شوند، می‌تواند به‌طور نامحدودی ادامه پیدا کند، اما انرژی صوتی به‌وسیله دیواره‌های نرم و اینرسی هوا جذب می‌شود؛ همان‌طور که اصطکاک، انرژی تاب را جذب می‌کند. بنابراین، وقتی به بازخوان‌های دستگاه گفتار (سازه‌ها) نگاه می‌کنیم، مقداری پهنای باند دارند، به این علت که فرکانس‌های بازخوانی دستگاه گفتار میرا هستند. اگر دیواره‌های دستگاه گفتار سخت بودند (و در نتیجه می‌توانستند انرژی صوتی را بی‌آنکه جذب کنند، بازتاب دهند)، سازه‌ها، پهنای باند بسیار کوچک‌تری داشتند.



شکل ۹-۲. طیف قدرت شکل موج‌هایی که در شکل ۹-۱ نشان داده شد. توجه کنید که افزایش میرایی متناظر با افزایش پهنای باند قله طیف است.

پهنای باند سازه‌ها در آواهای خیشومی، پهن‌تر از آواهای غیرخیشومی است، زیرا مساحت سطح^۱ و حجم دستگاه گفتار وقتی بینی باز است، بیشتر می‌شود. مساحت سطح بزرگ‌تر یعنی اینکه دیواره‌های دستگاه گفتار در مقایسه با آواهای غیرخیشومی، صوت بیشتری را جذب می‌کند و حجم بیشتر هوا یعنی اینکه اینرسی هوای درون دستگاه گفتار نیز صوت بیشتری را جذب می‌کند. با این وجود، همان‌طور که بعداً خواهیم دید، گسترش پهنای باند F_1 در واژه‌های خیشومی بسیار پیچیده‌تر از این است.

۹-۲. انسدادی‌های خیشومی

با خیشومی ملازی [N] شروع می‌کنیم، زیرا ساده‌ترین همخوان خیشومی برای توصیف است (تحلیلی که من در اینجا ارائه می‌دهم، تا حدودی در ادامه توصیف فنی‌تر فوجیمورا^۲، ۱۹۶۲ است. لازم به ذکر است که لینک‌ویست-گوفین^۳ و ساندبرگ^۴، ۱۹۷۶ اظهار کرده‌اند که حفره‌های سینوسی نقش مهمی در آکوستیک

1. surface area

2. Fujimura

3. Lindqvist-Gauffin

4. Sundberg

خیشومی‌ها ایفا می‌کنند. فاکتوری که در حاشیه تحلیل مورد بحث قرار می‌گیرد). وقتی ملاز پائین می‌آید و بدنه زبان برای تولید همخوان خیشومی ملازی بالا می‌رود، در نگاه اول می‌توان دستگاه گفتار را به صورت لوله واحدی که در چاکنا بسته و در سوراخ‌های بینی باز است، در نظر گرفت. اگر طول لوله را بدانیم، می‌توانیم فرکانس‌های بازخوانی آن را محاسبه کنیم، زیرا این، یک بازخوان یک‌چهارم طول موج است (مانند شکل‌بندی دستگاه گفتار برای تولید شوا، فصل ۲). در تصاویر اشعه ایکس که توسط فانت (۱۹۶۰) منتشر شده، فاصله بین ملاز تا سوراخ‌های بینی ۱۲/۵ سانتی‌متر و فاصله چاکنا تا ملاز ۹ سانتی‌متر است و بنابراین، مجموع فاصله چاکنا تا سوراخ‌های بینی ۲۱/۵ سانتی‌متر است. در این صورت، چهار فرکانس بازخوانی لوله عبارت‌اند از (c سرعت صوت در هوا بر حسب سانتی‌متر بر ثانیه):

$$F_1 = c/4L = 35000/4 \times 21/5 = 35000/86 = 407 \text{ Hz}$$

$$F_2 = 3c/4L = 1221 \text{ Hz}$$

$$F_3 = 5c/4L = 2035 \text{ Hz}$$

$$F_4 = 7c/4L = 2849 \text{ Hz}$$

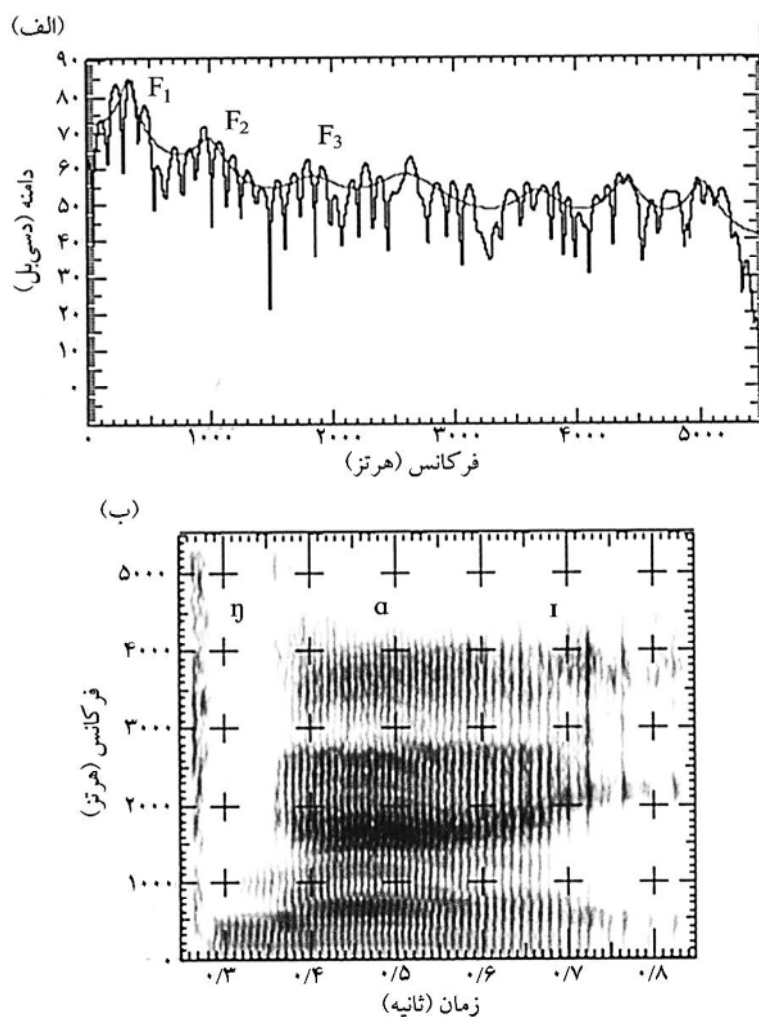
این مقادیر برای سازه‌های خیشومی [N] از جهتی غیردقیق هستند، زیرا فرض ما که می‌توان دستگاه گفتار را به صورت یک لوله واحد در نظر گرفت، نادرست است. اگر فرض کنیم که یک گرفتگی خیشومی در سوراخ‌های بینی وجود دارد (به عبارتی بینی «گردی لب» دائمی دارد)، نتایج را می‌توان بهبود بخشید. براساس نظریه اختلال، می‌دانیم که یک نقطه حداکثر سرعت برای هر یک از بازخوان‌ها در سوراخ‌های بینی وجود دارد، پس انتظار داریم که مقدار هر یک از فرکانس‌های بازخوانی کمتر از محاسبات بالا باشد. این تنها یکی از موارد تفاوت آناتومی مجرای بینی با یک لوله واحد است. پیش‌بینی کمی فرکانس‌های سازه‌ای [N] دشوار است، زیرا شکل مجرای بینی از شخصی به شخص دیگر متفاوت است - و برای هر شخصی از یک روز به روز دیگر متفاوت است، مثلاً وقتی که سرما خورده‌اید - اما در عین حال، یک ویژگی اصلی قابل تشخیص است. سازه‌ها در خیشومی ملازی فاصله کمتری از یکدیگر دارند تا در شوا. در یک دستگاه گفتار مردانه، فاصله بین سازه‌ها برای واکه شوا

در حدود ۱۰۰۰ هرتز است، در حالی که برای [N] در حدود ۸۰۰ هرتز است. شکل ۳-۹ (الف)، طیف [ŋ] را در کلمه تایلندی [ŋâ:t] به معنی «آسان» نشان می‌دهد. این طیف، همان‌طور که پیش‌بینی شد، چهار سازه با فاصله تقریباً مساوی در زیر ۳ کیلوهرتز دارد و فرکانس آنها تقریباً معادل فرکانس‌هایی است که برای یک لوله واحد با طول ۲۱/۵ سانتی‌متر بدست آوردیم.

پایین‌بودن دامنه F_3 ممکن است به دلیل حفره‌های سینوسی پیشین باشد (بعداً در این مورد بیشتر بحث خواهد شد). شکل ۳-۹ (ب) طیف‌نگاشت همان هجایی است که طیف‌های شکل ۳-۹ (الف) از آن گرفته شده‌اند. در طیف‌نگاشت می‌توانید ببینید که فرکانس F_1 و F_2 در انسدادی خیشومی [ŋ] پایین‌تر از واکه [a] است. در طیف‌نگاشت خیشومی یک فضای خالی در حدود ۲۰۰۰ هرتز نیز مشاهده می‌شود که متناظر با F_3 ضعیف در طیف‌هاست و سازه‌های بالای ۳ کیلوهرتز آنقدر ضعیف هستند که در طیف‌نگاشت ظاهر نمی‌شوند. همخوان‌های خیشومی عموماً ضعیف‌تر از واکه‌ها هستند (یعنی دامنه کمتری دارند). بخشی از آن به دلیل این واقعیت است که مجرای بازخوانی بزرگ‌تر موجب افزایش میرایی سازه‌ها می‌شود. اما دلیل اصلی اینکه چرا همخوان‌های خیشومی ضعیف‌تر از واکه‌ها هستند، این است که (۱) در تولید خیشومی‌ها، دستگاه گفتار شامل **حفره‌های جانبی** مانند سینوس‌هاست و (۲) میزان گرفتگی دستگاه گفتار در همخوان‌های خیشومی بیشتر از واکه‌ها می‌باشد (گرفتگی در ورودی مجرای بینی قرار دارد).

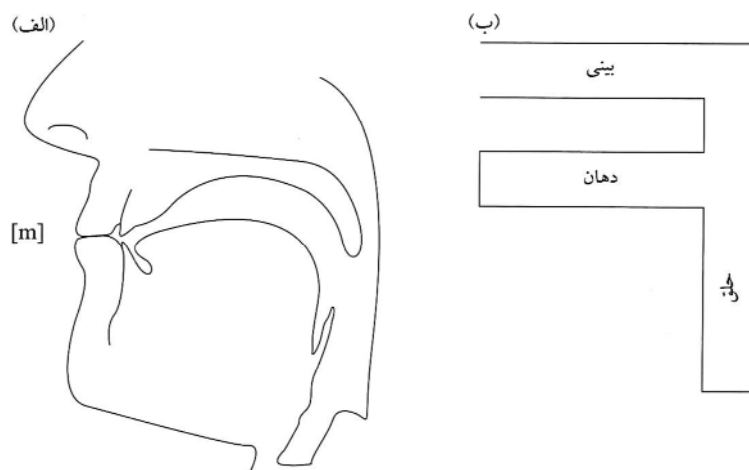
تفاوت اصلی بین [N] و [m] در این است که حفره دهان به صورت یک **شاخه جانبی** در لوله بازخوانی ایجاد می‌شود. ردیابی اشعه X در شکل ۴-۹ (الف)، حفره دهان را نشان می‌دهد و مدل لوله‌ای شکل ۴-۹ (ب)، دهان را به صورت یک شاخه جانبی برای لوله‌ای که حلق و بینی می‌سازند، نشان می‌دهد. حفره دهان را می‌توان به صورت لوله‌ای که در یک انتها بسته (لب‌ها) و در انتهای دیگر باز (ملاز) است و با طول ۸ سانتی‌متر مدل‌سازی نمود. بدین ترتیب، می‌توانیم بازخوان‌های حفره دهان را محاسبه کنیم، همان‌طور که برای شوا (در فصل ۵) یا برای [N] (در پاراگراف‌های قبلی) محاسبه کردیم. با فرض اینکه حفره دهان یک لوله واحد با طول ۹ سانتی‌متر

است که در یک انتها باز و در انتهای دیگر بسته است، پایین‌ترین فرکانس بازخوانی آن ۹۷۲ هرتز $(9 \times 4) / 35000 = 972 \text{ Hz}$ و دومین فرکانس بازخوانی آن ۲۹۱۷ هرتز $(3 \times 972 = 2916 \text{ Hz})$ است.



شکل ۹-۳. (الف) طیف‌های اف‌اف‌تی و ال‌پی‌سی زمزمه خیشومی^۱ در [ŋ] که توسط یک گویشور مرد تایلندی تولید شده است. طیف این همخوان زیر ۳۰۰۰ هرتز چهار سازه دارد. (ب) طیف‌نگاشت واژه‌ای که طیف‌ها از آن گرفته شده‌اند. طیف‌ها از نقطه میانی زمزمه خیشومی، در زمان حدوداً ۰/۳۳ ثانیه گرفته شده‌اند.

فرکانس‌های بازخوانی حفره دهان در خیشومی‌ها شبیه فرکانس‌های بازخوانی که قبلاً دیدیم نیستند، زیرا در این همخوان‌ها، حفره دهان یک شاخه جانبی برای یک لوله بازخوانی بزرگ‌تر است و مستقیماً به فضای بیرون باز نمی‌شود. از این رو، مؤلفه‌های فرکانسی منبع صوت که حول وحوش فرکانس‌های بازخوانی حفره جانبی هستند، در شاخه جانبی بازخوان می‌شوند، بی‌آنکه در برون‌داد آکوستیکی سیستم لوله تظاهری داشته باشند. در واقع، این مؤلفه‌ها در شاخه جانبی «جذب می‌شوند». بنابراین، مؤلفه‌های فرکانسی خیشومی [m] که حول وحوش فرکانس‌های بازخوانی دهان هستند، خنثی شده و در برون‌داد آکوستیکی تبدیل به پادبازخوان^۱ (یا همان پادسازه) می‌شوند. سازه‌ها در طیف به صورت قله‌های انرژی صوتی ظاهر می‌شوند و پادسازه‌ها به صورت دره‌های برجسته طیفی. بسیار مهم است که بین عدم وجود بازخوان با تولید فعالانه پادبازخوان تمایز قائل شویم. در طیف شوآ دره‌هایی وجود دارد که ناشی از عدم وجود بازخوان هستند، صرفاً به این دلیل که برخی از مؤلفه‌های فرکانسی به اندازه بقیه تقویت نشده‌اند. در مقابل این، مؤلفه‌های فرکانسی در [m] وجود دارد که فعالانه از طیف کسر می‌شوند.



شکل ۹-۴. (الف) وضعیت اندام‌ها هنگام تولید [m] با ردیابی اشعه X (اقتباس از استراکا^۲، ۱۹۶۵، صفحه ۳۴). (ب) یک مدل لوله‌ای از این شکل‌بندی دستگاه گفتار.

1. anti-resonance

2. Straka

تأثیر دیگر پادسازه بر طیف این است که دامنه همه سازه‌هایی که بالاتر از پادسازه هستند، کاهش پیدا می‌کند (محاسبه «فانت» نشان می‌دهد که دامنه همه سازه‌های بالاتر از پادسازه در حدود $1/6$ دسی‌بل در اکتاو کاهش پیدا می‌کند)، زیرا این سازه‌ها بر حاشیه پادسازه سوار می‌شوند. اسامو فوجیمورا به من اظهار کرد (طی یک گفتگوی شخصی) که کاهش دامنه یک ویژگی کلی پادسازه‌ها نیست، بلکه به میزان تفکیک فرکانسی بازخوان و پادبازخوان که با یک حفره جانبی به هم مرتبط شده‌اند، بستگی دارد. در بحث پادسازه‌ها که در اینجا آمده، از پرداختن به برخی جزئیات درباره اینکه چگونه پادسازه‌ها از حفره‌های جانبی منتج می‌شوند، صرف نظر شده است (به‌عنوان مثال، ارتباط آنها با جفت‌شدگی آکوستیکی و این واقعیت که پادبازخوان‌ها همواره با بازخوان‌ها جفت می‌شوند). و بنابراین، برآورد کاهش دامنه به‌میزان $1/6$ دسی‌بل در اکتاو، بدون اثبات به‌عنوان یک برآورد کلی مناسب که تأثیر طیفی یک پادسازه را نشان می‌دهد، پذیرفته شد. بنابراین، آواهای خیشومی انرژی بیشتری در انتهای پایینی طیف دارند. افزایش پهنای باند که در آواهای خیشومی مشاهده می‌شود، باعث کاهش دامنه انرژی سازه‌ای نیز می‌شود. به‌همین دلیل آواهای خیشومی در طیف‌نگاشت، ضعیف‌تر از واکه‌های دهانی مجاور ظاهر می‌شوند.

پادسازه و پادماده^۱

هرگاه مبحث پادسازه را در کلاس مطرح می‌کنم، یکی از من می‌پرسد که آیا پادسازه شبیه پادماده است. پاسخ این است «من نمی‌دانم، شاید». اگر ماده و پادماده یکدیگر را خنثی کنند، همان‌طور که در سفینه اینترپرایز^۲ می‌کنند، منطقی به نظر می‌رسد اگر بگوییم پادسازه کمی شبیه پادماده است. لرد ریلی^۳ (۱۸۹۶) ویژگی خنثی‌کنندگی انرژی صوتی در پادسازه‌ها را ناشی از «تداخل امواج مستقیم و امواج انعکاسی» می‌داند. در انسدادی‌های خیشومی، امواج مستقیم امواج صوتی هستند که در لوله حلق-بینی

←

1. anti-matter

۲. Enterprise. از سری فیلم‌های تخیلی. (مترجم)

3. Lord Rayleigh

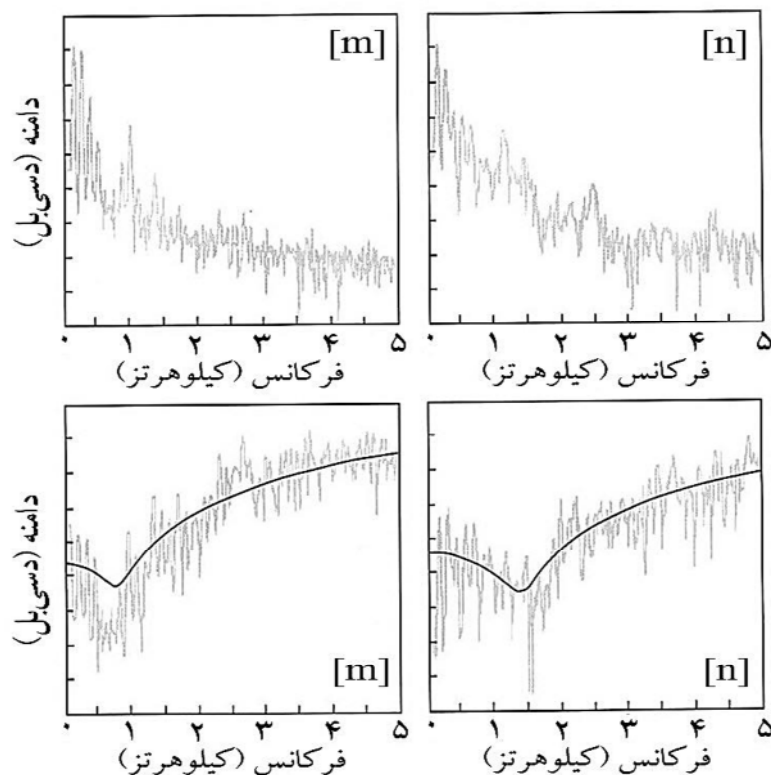
حرکت می‌کنند و امواج انعکاسی امواجی هستند که در حفره دهان بازخوان می‌شوند. به دلیل تغییر قطبیتی که در انتهای باز لوله‌ای که در یک انتها بسته است، روی می‌دهد (مثال شلاق گاو در فصل ۲)، انعکاس بازخوان‌های حفره دهان، دقیقاً در فاز مخالف مؤلفه‌های فرکانسی مشابه در لوله حلق-بینی است. بنابراین، وقتی یک موج مثبت است، موج دیگر منفی است و وقتی شما این دو را به هم اضافه می‌کنید، یکدیگر را خنثی می‌کنند. بر این اساس، فکر می‌کنم پادسازه شبیه پادماده (هرچه که باشد) است.

فرکانس برخی پادسازه‌ها در طیف یک انسدادی خیشومی به طول حفره دهان بستگی دارد. به عنوان مثال، طول حفره دهان در [n] در حدود ۵/۵ سانتی‌متر است. فرکانس‌های بازخوانی لوله واحدی که در یک انتها بسته و در انتهای دیگر باز است و طول آن ۵/۵ سانتی‌متر است، در ۱۵۹۱ هرتز و ۴۷۷۳ هرتز است. بنابراین، انتظار می‌رود طیف [n] یک پادسازه در حدود ۱۶۰۰ هرتز و یک پادسازه دیگر در حدود ۴۸۰۰ هرتز داشته باشد. همان‌طور که در بالا گفته شد، طول حفره دهان در [m] در حدود ۹ سانتی‌متر است، بنابراین انتظار داریم در طیف [m] یک پادسازه در کمتر از ۱۰۰۰ هرتز و یک پادسازه دیگر در حدود ۳۰۰۰ هرتز ببینیم. بدین ترتیب، فرکانس پادسازه‌ها سرنخ‌هایی برای محل تولید خیشومی‌های مذکور محسوب می‌شود.

شکل ۵-۹ نتایج یک روش محاسبه فرکانس پادسازه‌ها در همخوان‌های خیشومی را نشان می‌دهد (مارپل^۱، ۱۹۸۷؛ کی^۲، ۱۹۸۹). در این روش، که تحلیل میانگین متحرک خودبازگشتی^۳ نامیده می‌شود، ابتدا با تجزیه ال‌پی‌سی، قله‌های طیف خیشومی مشخص و سپس سیگنال با هدف حذف سازه‌ها فیلتر می‌شود و در نهایت، تجزیه ال‌پی‌سی برای پیدا کردن پادسازه‌ها به کار می‌رود. نمودارهای بالای شکل، طیف [m] و [n] را که توسط یک گویشور مرد انگلیسی‌زبان تولید شده‌اند، نشان می‌دهد و در نمودارهای پایین، طیف سیگنال‌ها پس از حذف سازه‌ها با روش «فیلتر معکوس»^۴، به همراه طیف ال‌پی‌سی هر همخوان که پایین‌ترین پادسازه را نشان می‌دهد، قابل مشاهده است. این شکل نشان می‌دهد که فرکانس محاسبه‌شده اولین

1. Marple 2. Qi 3. autoregressive moving average analysis 4. inverse filtering

پادسازه در $[m]$ و $[n]$ اندکی پایین‌تر از مقدار پیش‌بینی شده است. در $[m]$ اولین پادسازه (A_1) در حدود ۷۵۰ هرتز است، در حالی که طبق پیش‌بینی ما باید در حدود ۹۷۰ هرتز باشد و در $[n]$ مقدار A_1 در حدود ۱۴۰۰ هرتز است، در حالی که مقدار پیش‌بینی شده ۱۶۰۰ هرتز بود. بنابراین، مدل لوله‌ای ساده‌ای که به کار می‌بریم خیلی درست نیست، اما الگوی کلی که مدل پیش‌بینی می‌کند، مورد تأیید است. مقدار A_1 در $[m]$ پایین‌تر از $[n]$ است.



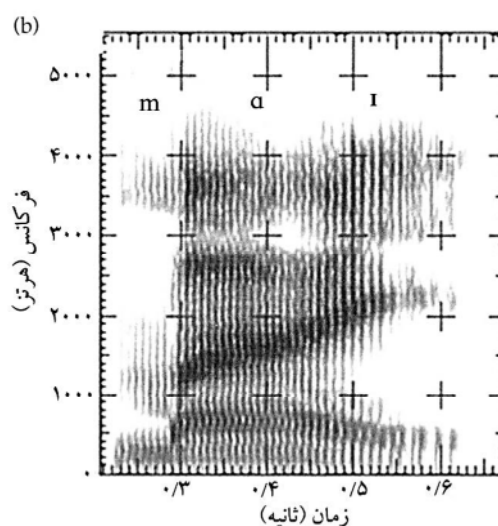
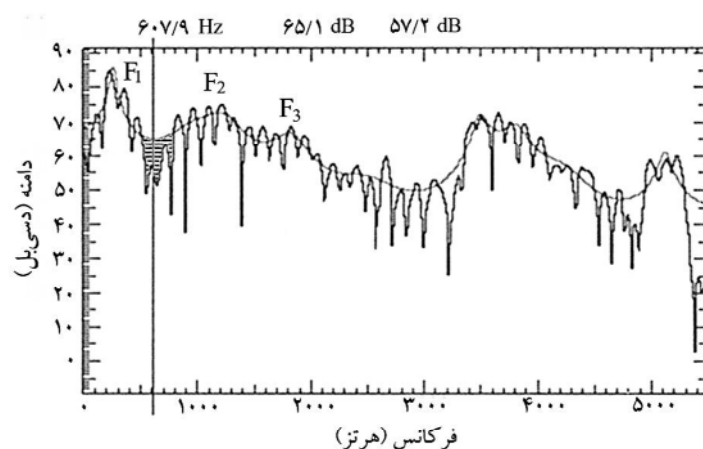
شکل ۹-۵. پادسازه‌های $[m]$ و $[n]$ در زبان انگلیسی. نمودارهای سمت چپ مربوط به $[m]$ و نمودارهای سمت راست مربوط به $[n]$ است. نمودارهای بالا، طیف اف‌ا‌ف‌تی خیشومی‌ها و نمودارهای پایین، طیف ال‌پی‌سی و طیف‌های فیلتر معکوس شده به همراه تأثیر فیلتر پادسازه‌ها را نشان می‌دهند. فاصله بین علامت‌گذاری‌ها روی محور عمودی ۱۰ دسی بل است. (اقتباس از کی ۱۹۸۹، شکل‌های ۲۷-۲۴) و انتشار با کسب اجازه)

حال، با دانستن اینکه چگونه پادسازه‌ها به وسیله حفره‌های جانبی در دستگاه گفتار به وجود می‌آیند، اهمیت سینوس‌های بینی روشن‌تر می‌شود. سینوس‌ها شبیه بازخوان‌های «هلمهولتز» عمل می‌کنند. بنابراین، فرکانس‌های بازخوانی آنها یا به عبارتی، فرکانس پادسازه‌هایی که آنها وارد طیف می‌کنند، به حجم سینوس و ابعاد منفذ ورودی آن بستگی دارد. تنوعات بسیاری به لحاظ اندازه سینوس و اندازه منفذ ورودی به سینوس بین افراد مشاهده می‌شود. بنابراین، به سختی می‌توان برآوردی قطعی از فرکانس پادسازه‌های حاصل از سینوس‌ها به دست داد. مطالعه لینک‌ویست-گوفین و ساندبرگ (۱۹۷۶) در یک گویشور نشان می‌دهد که سینوس فکی، پادسازه‌ای در حدود ۵۰۰ هرتز و سینوس پیشانی، پادسازه‌ای در حدود ۱۴۰۰ هرتز ایجاد می‌کند.

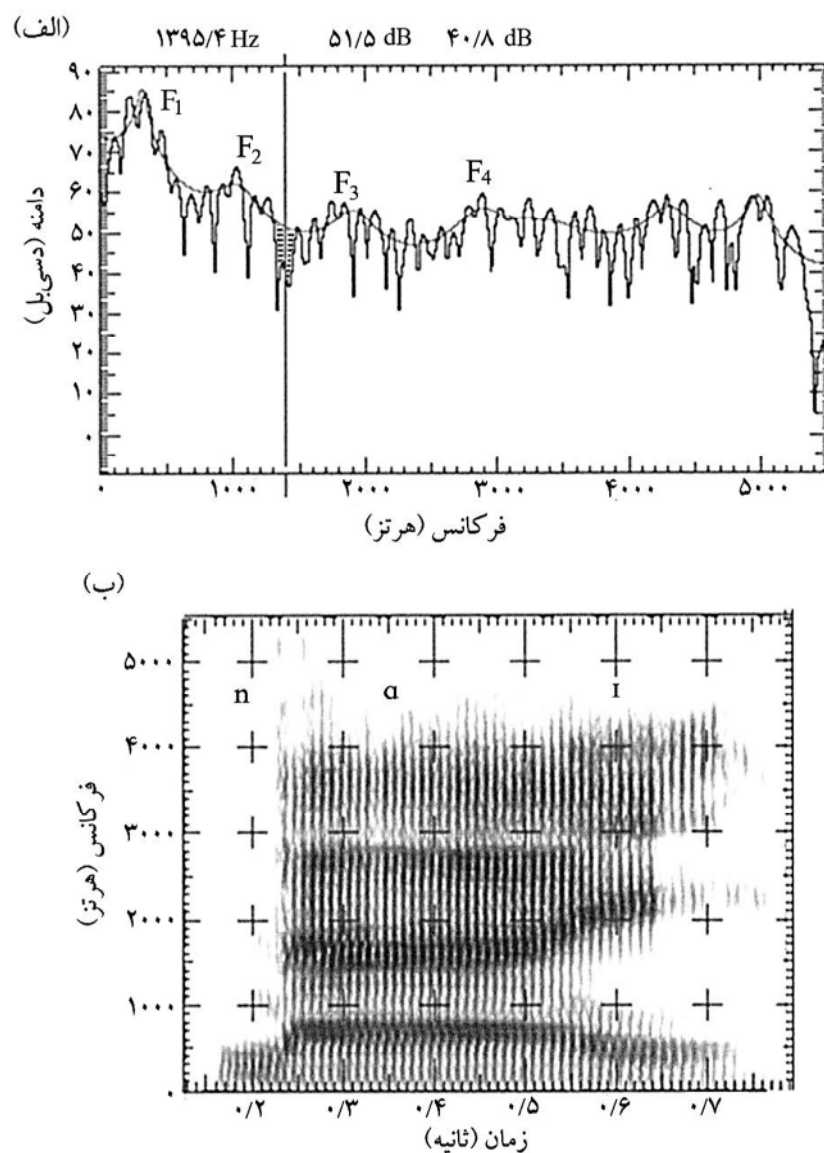
شکل‌های ۹-۶ و ۹-۷، طیف افاف‌تی و ال‌پی‌سی و نیز طیف‌نگاشت [m] و [n] در کلمات [mâi] «نه» و [na:i] «رئیس» که توسط یک گویشور مرد تایلندی تولید شده‌اند را نشان می‌دهد. پادسازه اول در طیف‌ها با هاشور و مکان‌نمای عمودی نشان داده شده است. اگر به طیف‌ها نگاه کنید، می‌بینید که پادسازه‌ها نوعی عدم تطابق بین طیف ال‌پی‌سی با طیف فوریه ایجاد می‌کنند. دلیل آن، این است که در تجزیه ال‌پی‌سی فرض بر این است که تابع فیلتر دستگاه گفتار هیچ پادسازه‌ای ندارد. گفتیم که حفره دهان در [m] به لحاظ نظری باید بازخوان‌هایی در حدود ۹۷۲ هرتز و ۳۰۰۰ هرتز داشته باشد که در طیف [m] به صورت پادسازه ظاهر می‌شوند، در حالی که طیف [n] باید پادسازه‌هایی در فرکانس‌های بالاتر (۱۶۰۰ هرتز و ۴۸۰۰ هرتز) داشته باشد. این پیش‌بینی‌ها به طور کلی در طیف‌های شکل ۹-۶ تأیید می‌شوند. توجه داشته باشید که پادسازه می‌تواند در طیف‌نگاشت به صورت یک نوار سفید ظاهر بشود یا نشود. اگر هیچ سازه‌ای نزدیک به پادسازه تولید نشود، پادسازه به صورت یک نوار سفید در طیف‌نگاشت ظاهر می‌شود (شکل ۹-۶ (ب)) اما اگر فرکانس پادسازه تقریباً برابر با فرکانس یک سازه باشد، نتیجه نهایی آن تضعیف قله سازه است.

توصیف [m] و [n] نیز مانند [N] از جهتی دقیق نیست، زیرا فرض ما در مورد

مدل لوله‌ای ساده‌انگارانه است. با این وجود، در این تحلیل ویژگی‌های اصلی خیشومی‌ها توصیف می‌شود: F_1 پایین (که گاهی «سازه خیشومی» نامیده می‌شود)، فاصله کم بین سازه‌ها و وجود پادسازه‌ها که فرکانس آنها بر اساس محل تولید همخوان انسدادی به دست می‌آید.



شکل ۹-۶. (الف) طیف‌های اف اف تی و ال پی سی زمزمه خیشومی در [m] که توسط یک گویشور مرد تایلندی تولید شده است. برآورد فرکانس اولین پادسازه با مکان‌نمای عمودی و هاشور نشان داده شده است. (ب) طیف‌نگاشت واژه‌ای که طیف‌ها از آن گرفته شده‌اند. طیف‌ها از نقطه میانی زمزمه خیشومی در زمان حدود ۰/۲۶ ثانیه گرفته شده‌اند.



شکل ۹-۷. (الف) طیف‌های اف اف تی و ال پی سی زمزمه خیشومی در [n] که توسط یک گویشور مرد تایلندی تولید شده است. برآورد فرکانس اولین پادسازه با مکان‌نمای عمودی و هاشور نشان داده شده است. (ب) طیف‌نگاشت واژه‌ای که طیف‌ها از آن گرفته شده‌اند. طیف‌ها از نقطه میانی زمزمه خیشومی در زمان حدود ۰/۲ ثانیه گرفته شده‌اند.

درک پادسازه‌ها

آزمایش رپ^۱ (۱۹۸۶) نشان می‌دهد که شنوندگان در ۷۲ درصد موارد می‌توانند زمزمه‌های خیشومی استخراجی از هجاهای آغازشده با [n] و [m] را درست شناسایی کنند. با افزودن گذرهای سازه‌ای واکه (فصل ۸)، شناسایی درست به بیش از ۹۵ درصد و با افزودن فقط ۱۰ میلی‌ثانیه از واکه بعدی، به ۹۴ درصد رسید. این نتایج، حاکی از این است که فرکانس پادسازه‌ها در [m] و [n] سرخ‌های درکی مهمی برای این همخوان‌ها محسوب می‌شوند، یا شاید هم نمی‌شوند. هفتادودو درصد پاسخ درست در یک وظیفه شناسایی، در حالی که صرفاً با حدس ۵۰ درصد پاسخ‌ها درست خواهند بود، یعنی اینکه زمزمه خیشومی خیلی هم جنبه اطلاعی نداشته است. در واقع، برای به‌دست‌آوردن این نتیجه، فقط ۱۲/۵ درصد از آواها باید نوعی سرخ برای تشخیص محل تولید داشته باشند.

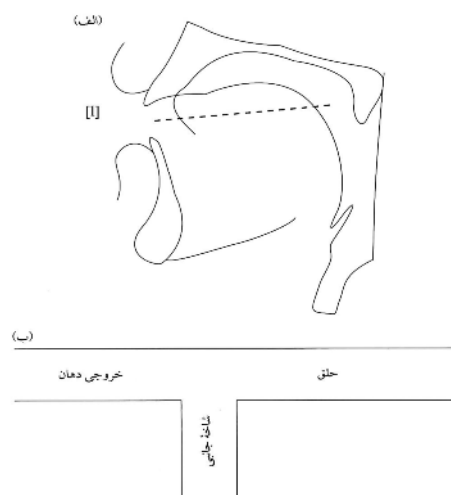
یکی از دلایل اینکه چرا زمزمه‌های خیشومی ممکن است سرخ‌های خیلی خوبی برای محل تولید نباشند، این است که در تجربه معمولی (برخلاف بسیاری از آزمایش‌های درکی-گفتاری آزمایشگاهی) ارتباط گفتاری در میان صداهای پس‌زمینه رخ می‌دهد. بنابراین، ممکن است پادسازه‌ها (یا به‌طور دقیق‌تر، تغییراتی که در شکل طیف به‌واسطه پادسازه‌ها ایجاد می‌شود) در سیگنال خیلی قابل شنیدن نباشند، زیرا صداهای پس‌زمینه، دره‌های پادسازه را پر می‌کنند. همچنین، ممکن است شنوندگان توجه چندانی به جزئیات شکل طیف در زمزمه‌های انسدادی خیشومی نکنند، زیرا طیف حاوی اطلاعات شخصی درباره اندازه و شکل (فعلی) سینوس‌های گویشور است که احتساب آنها به‌عنوان یک سرخ برای محل تولید کاملاً بی‌ربط می‌باشد چراکه در بین گویشوران خیلی قابل پیش‌بینی نیست.

۹-۳. همخوان‌های کناری

همخوان‌های کناری نیز با یک شاخه جانبی که سبب ایجاد پادسازه در طیف خروجی می‌شود، تولید می‌شوند. پس، می‌توانیم به تحلیل ویژگی‌های آکوستیکی آنها شبیه به تحلیل خیشومی‌ها در بخش ۹-۲ پردازیم. شکل ۹-۸ (الف) تولید [I] را با ردیابی اشعه X نشان می‌دهد. استدلال فانت (۱۹۶۰) این است که هنگام تولید

1. Repp

همخوان‌های کناری، یک کیسه کوچک هوا بالای زبان ایجاد می‌شود که برای مجرای آکوستیکی اصلی که حول یک یا هر دو کناره زبان انحنای پیدا می‌کند، مانند یک شاخه جانبی عمل می‌کند. در تصاویر گرفته‌شده از شکل زبان با امواج فراصوت^۱ (استون^۲، ۱۹۹۱)، نوعی حرکت موج‌مانند^۳ از یک سو به سوی دیگر زبان هنگام تولید [I] در انگلیسی آمریکایی مشاهده می‌شود که نشان می‌دهد یک کناره زبان بازتر از کناره دیگر آن است. مشاهده کام‌نگاشت^۴ کناری‌ها در زبان‌های تامیل و تایوانی نشان می‌دهد که هر دو کناره زبان تا حدی پایین می‌آید، اما ممکن است نوعی عدم تقارن بین کناره‌های چپ و راست زبان به لحاظ میزان بازبودن آنها وجود داشته باشد. به هر حال، یک کیسه هوا بر روی زبان برجا می‌ماند و به‌عنوان یک شاخه جانبی تولید پادسازه عمل می‌کند، همان‌طور که حفره دهان در همخوان‌های خیشومی تولید می‌کند.



شکل ۸-۹. (الف) تولید [I] در ایسلندی با ردیابی اشعه X. نقطه چین نشان‌دهنده این است که حداقل یکی از کناره‌های زبان باز است. اقتباس از پتورسون^۵ (۱۹۷۳، ص ۹۱). (ب) یک مدل لوله‌ای ساده از این شکل‌بندی دستگاه گفتار. مجرای جانبی به وسیله یک کیسه هوا بر روی زبان شکل می‌گیرد، در صورتی که مجرای خروجی دور یک یا هر دو کناره زبان تشکیل می‌شود.

این شکل‌بندی دستگاه گفتار را می‌توان به صورت یک لوله واحد با یک شاخه جانبی کوتاه که همان کیسه هوا بر روی زبان است، مدل‌سازی کرد. این مدل لوله‌ای در شکل ۸-۹ (ب) نشان داده شده است. محاسبه فانت (۱۹۶۰) نشان می‌دهد که

1. ultrasound 2. Stone 3. rocking motion 4. palatogram 5. Petursson

برای یک گویشور، طول کیسه هوا ۴ سانتی‌متر و طول دستگاه گفتار ۱۶ سانتی‌متر است (۱۰ سانتی‌متر از چاکنا تا شاخه جانبی و یک حفره خروجی کناری به طول ۶ سانتی‌متر). بنابراین، مقدار بازخوان‌های دستگاه گفتار ۵۳۱ هرتز، ۱۵۹۴ هرتز، و ۲۶۵۶ هرتز است (البته اگر خلاف واقع فرض کنیم که یک لوله واحد داریم). شما ممکن است این فرکانس‌ها را به عنوان فرکانس‌های بازخوانی شوآ بشناسید. کیسه هوا که به صورت لوله‌ای در یک انتها باز و در انتهای دیگر بسته مدل‌سازی شد، در ۲۱۲۵ هرتز بازخوان می‌شود ($c/4L$ ، که در آن $c = ۳۵۰۰۰$ cm و $L = ۴$ cm است). از آنجا که کیسه هوا یک حفره جانبی است، بازخوان مذکور در لب‌ها تبدیل به یک پادبازخوان در برون‌داد می‌شود.

این مدل لوله‌ای برای همخوان‌های کناری تقریباً بیش از حد ساده‌انگارانه است. در واقع، قطر حفره خروجی نسبت به لوله‌ای که از چاکنا تا گرفتگی کناری امتداد دارد، کوچک‌تر است. مهم‌ترین پیامد آکوستیکی آن، این است که فرکانس سازه اول کمتر از آن مقداری است که در یک لوله واحد خواهد بود. کناری [l] در پایان هجا در انگلیسی آمریکایی نیز با گرفتگی بدنه زبان که فرکانس سازه دوم را پایین می‌آورد، تولید می‌شود. در عین حال، پیش‌بینی اصلی مدل لوله‌ای این است که علامت طیفی تولید کناری، وجود یک پادسازه حول و حوش فرکانس F_3 است.

شکل ۹-۹ طیف اف‌اف‌تی و ال‌پی‌سی همخوان [l] را نشان می‌دهد. این همخوان توسط همان گویشور تایلندی تولید شده است که آواهای دو شکل ۹-۳ و ۹-۶ را تولید کرده است. طیف‌ها از مرکز همخوان کناری در کلمه [la:i] به معنی «روباه‌ها» گرفته شده‌اند. همان‌طور که هاشور و مکان‌نمای عمودی نشان می‌دهد یک پادسازه درست بالای F_3 وجود دارد. نکته دیگر اینکه، سازه‌ها در این شکل نسبت به همخوان‌های خیشومی که پیش‌تر نشان دادیم، فاصله نسبتاً بیشتری از یکدیگر دارند. در اینجا فقط سه قله سازه‌ای زیر ۴ کیلوهرتز تولید شده است، در حالی که در آواهای خیشومی چهار سازه وجود داشت. ممکن است یک پادسازه دیگر نیز حدود ۱ کیلوهرتز در طیف قابل مشاهده باشد که احتمالاً ناشی از عدم تقارن در میزان بازبودن گرفتگی کناری در اطراف زبان است.

در همخوان‌های کناری نیز مانند خیشومی‌ها، حضور یک پادسازه سبب کاهش دامنه همه سازه‌های بالاتر در حدود ۱/۶ دسی‌بل می‌شود. به علاوه، باریک‌بودن حفره خروجی باعث می‌شود که همخوان‌های کناری عموماً نرم‌تر از واکه‌ها تولید شوند. این جنبه‌های آکوستیکی همخوان‌های کناری را می‌توانید در طیف‌نگاشت شکل ۹-۹ (ب) ببینید. یکی از تفاوت‌های مهم بین کناری‌ها و خیشومی‌ها این است که میانگین فاصله بین سازه‌ها در کناری‌ها بیشتر از خیشومی‌هاست، به این دلیل که لوله بازخوانی اصلی در خیشومی‌ها بلندتر از کناری‌هاست. در مثال‌هایی که به آنها اشاره شد، میانگین فاصله بین سازه‌ها در خیشومی‌ها در حدود ۸۰۰ هرتز است، در حالی که در کناری‌ها ۱۰۰۰ هرتز است.

۹-۴. خیشومی‌شدگی^۱

حال به بحث در مورد پیچیده‌ترین شکل‌بندی دستگاه گفتار می‌پردازیم (به فوجیمورا، ۱۹۶۲ و مائدا^۲، ۱۹۹۳ نیز مراجعه کنید). در واکه‌های خیشومی‌شده^۳ دو سیستم بازخوانی هم‌زمان عمل می‌کند: یکی تشکیل شده از حفره حلق به‌اضافه حفره دهان، و دیگری تشکیل شده از حفره حلق به‌اضافه حفره خیشوم. این بازخوان‌ها در شکل ۹-۱۰ نشان داده شده‌اند. در بحث‌های پیشین فرکانس‌های بازخوانی این دو سیستم را به دست آوردیم. فرکانس‌های بازخوانی سیستم حلق به‌اضافه دهان، یا مجرای دهانی^۴ که به‌صورت یک لوله واحد مدل‌سازی شد، در حدود ۵۰۰ هرتز، ۱۵۰۰ هرتز و ۲۵۰۰ هرتز است، اما فرکانس‌های بازخوانی سیستم حلق به‌اضافه بینی، یا مجرای خیشوم^۵ که به‌صورت یک لوله واحد مدل‌سازی شد، در حدود ۴۰۰ هرتز، ۱۲۰۰ هرتز و ۲۰۰۰ هرتز است. همه این سازه‌ها در طیف واکه‌های خیشومی‌شده حضور دارند. همان‌طور که دیدیم، فرکانس‌های بازخوانی مجرای دهانی با حرکات زبان و لب‌ها تغییر می‌کنند. بازخوان‌های مجرای خیشوم نیز با ایجاد تغییراتی در حفره حلق می‌توانند تغییر کنند و انتظار می‌رود که بازخوان‌های مجرای خیشوم به دلیل گرفتگی بینی در سوراخ‌های بینی، پایین‌تر از برآوردی باشد

1. nasalization

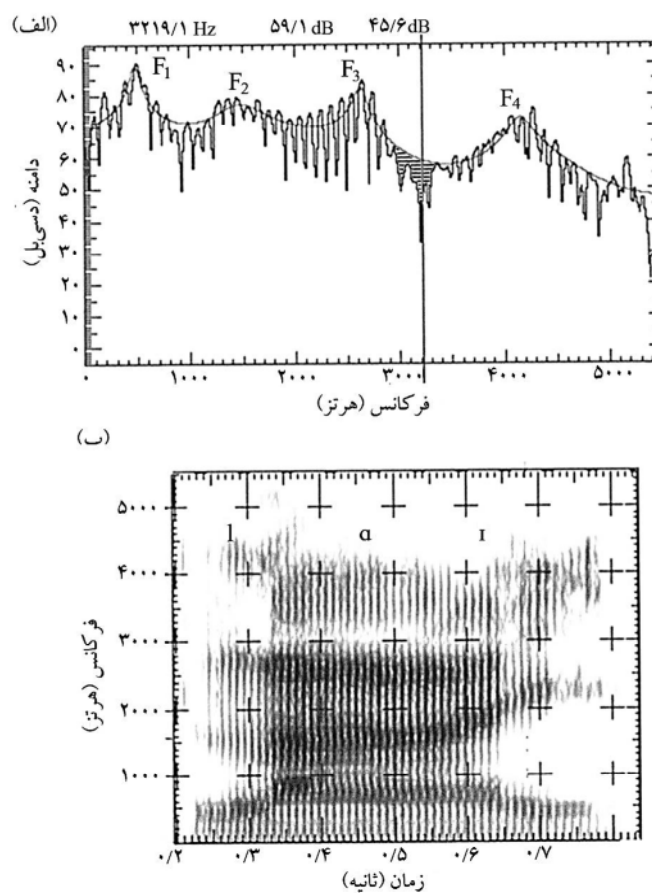
2. Maeda

3. nasalized vowels

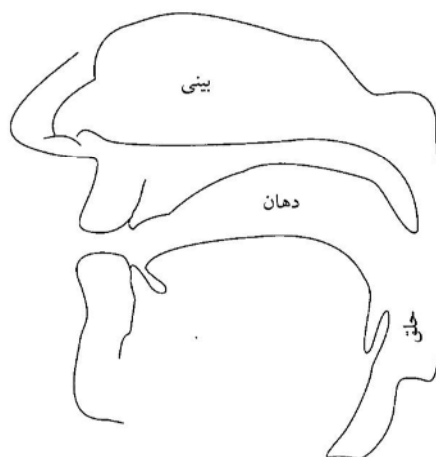
4. oral tract

5. nasal tract

که ما بر اساس مدل لوله واحد به دست می‌آوریم. با این‌همه، مدل‌های لوله‌ای پیش‌بینی می‌کنند که طیف یک واژه خیشومی، سازه‌های بسیاری خواهد داشت.



شکل ۹-۹. (الف) طیف اف‌اف‌تی و ال‌پی‌سی همخوان ناسوده کناری [l] که توسط یک گویشور تایلندی مرد تولید شده است. فرکانس پایین‌ترین پادسازه با مکان‌نمای عمودی و هاشور نشان داده شده است. (ب) طیف‌نگاشت کلمه‌ای که طیف‌ها از آن گرفته شده‌اند. طیف‌ها از نقطه وسط همخوان کناری در زمان ۰/۲۷ ثانیه گرفته شده‌اند.



شکل ۹-۱۰. شکل‌بندی دستگاه گفتار در واکه‌های خیشومی‌شده. یک سیستم آکوستیکی شامل حلق و دهان و سیستم آکوستیکی هم‌زمان دیگر شامل حلق و بینی است.

در بحث همخوان‌های خیشومی گفتیم که دهان بسته بازخوان‌هایی دارد که در طیف خیشومی‌ها تبدیل به پادبازخوان (پادسازه) می‌شوند. بازخوان‌های حفره بینی در واکه‌های خیشومی‌شده تبدیل به پادسازه می‌شوند؛ مانند پادسازه‌های همخوان‌های خیشومی، با این تفاوت که در واکه‌های خیشومی‌شده دهان بازتر از بینی است. در نتیجه، جفت‌شدگی آکوستیکی بین دهان با اتمسفر بیشتر از جفت‌شدگی بین مجرای بینی با اتمسفر است. فرکانس پادسازه‌ها در واکه‌های خیشومی‌شده تابعی از درجه جفت‌شدگی بین حفره بینی و حلق است. اگر جفت‌شدگی ضعیف باشد، فرکانس پادسازه تنها اندکی بیشتر از بازخوان‌های مجرای خیشوم است (۴۰۰ هرتز، ۱۲۰۰ هرتز، ۲۰۰۰ هرتز) اما در جفت‌شدگی آکوستیکی شدیدتر که در پیچه کامی حلقی^۱ کاملاً باز است، فرکانس پادسازه‌های خیشومی را می‌توان با فرض لوله‌ای که در یک انتها باز (در پیچه کامی حلقی) و در انتهای دیگر بسته (سوراخ‌های بینی) است، محاسبه کرد. از آنجا که فاصله ملاز تا سوراخ‌های بینی حدوداً $12/5$ سانتی‌متر است، دو پایین‌ترین فرکانس بازخوانی حفره بینی در حدود $680 (=c/4L)$

1. velopharyngeal port

و $2040 (3c/4L=)$ هرتز است. جدول ۹-۱ مقادیر پیش‌بینی‌شده این فرکانس‌های سازه‌ای و پادسازه‌ای را نشان می‌دهد.

اگر فقط نواحی زیر 1000 هرتز را در نظر بگیریم، واکه‌های غیرخیشومی تنها یک بازخوان (F_1) دارند، در صورتی که واکه‌های خیشومی شده سه برجستگی طیفی دارند: یک سازه دهانی (F_{10})، یک سازه خیشومی (F_{1n})، و یک پادسازه (A_1). فرکانس این مؤلفه‌ها به عوامل متعددی بستگی دارد. فرکانس F_{10} به موقعیت لب‌ها و زبان و فرکانس A_1 به درجه خیشومی‌شدگی بستگی دارد. با یک خیشومی‌شدگی جزئی، فرکانس A_1 تنها کمی بیشتر از F_{1n} است، اما با افزایش میزان جفت‌شدگی خیشومی، فرکانس A_1 افزایش پیدا می‌کند. پیامد آن، این است که A_1 در خیشومی‌شدگی جزئی، F_{1n} را خنثی کند و در درجات بالاتر خیشومی‌شدگی، F_{10} را. از این رو، به‌عنوان مثال، در طیف‌نگاشت یک $[a]$ با خیشومی‌شدگی شدید، ممکن است F_{1n} خیلی کم‌تر نسبت به مقدار معمول خود ظاهر شود، زیرا فرکانس F_{1n} پایین است (و همواره صرف نظر از نوع واکه پایین خواهد بود) و مقدار A_1 آن اندازه بالا هست که بتواند بیشتر انرژی F_{10} را خنثی کند.

جدول ۹-۱. فرکانس‌های پیش‌بینی‌شده برای سازه‌های دهانی و خیشومی و پادسازه‌های خیشومی در واکه‌های خیشومی‌شده (با فرض لوله واحد و با نادیده گرفتن تأثیرات جفت‌شدگی آکوستیکی).

	پادسازه‌ها	سازه‌های دهانی	سازه‌های خیشومی
	($L = 17/5$ cm)	($L = 17/5$ cm)	($L = 21/5$ cm)
F_1	۶۸۰	۵۰۰	۴۰۷
F_2	۲۰۴۰	۱۵۰۰	۱۲۲۱
F_3	—	۲۵۰۰	۲۰۳۵

این نوع تأثیر متقابل و پیچیده بین سازه‌های خیشومی، سازه‌های دهانی و پادسازه‌های خیشومی که در ناحیه فرکانسی F_1 روی می‌دهد، در سایر فرکانس‌ها نیز رخ می‌دهد. جای تعجب ندارد که تعداد واکه‌های خیشومی تمایزدهنده در زبان‌ها اغلب کمتر (و نه هرگز بیشتر) از تعداد واکه‌های دهانی است. هیچ مشکل گفتاری

برای تولید واکه‌های خیشومی شده وجود ندارد، اما این واکه‌ها به لحاظ آکوستیکی و درکی پیچیده‌تر از واکه‌های دهانی هستند.

یکی از پیامدهای تأثیر متقابل سازه‌ها و پادسازه‌ها در ناحیه F_1 واکه‌های خیشومی شده که در طیف‌نگاشت قابل مشاهده است، افزایش پهنای باند مؤثر F_1 است (زیرا در این ناحیه فرکانسی اغلب دو سازه وجود دارد، نه یک سازه). واکه شکل ۳-۹ (ب) برای من بسیار خیشومی شده به نظر می‌رسد و اگر شما این طیف‌نگاشت را با طیف‌نگاشت شکل ۹-۹ (ب) مقایسه کنید، می‌بینید که تشخیص F_1 در [la:t] آسان‌تر از تشخیص F_1 در واکه خیشومی شده است. این سازه‌ها که بسیار به هم نزدیک‌اند، شبیه به F_1 و F_2 واکه [u] هستند که در طیف‌نگاشت به نظر می‌رسد در یک سازه پهن ادغام شده‌اند. علاوه بر آن، همان‌طور که در مورد همخوان‌های کناری و خیشومی دیدیم، دامنه سازه‌ها به دلیل وجود یک پادسازه کاهش پیدا می‌کند. در آخر، می‌توانیم به برخی از الگوهای واجی اشاره کنیم که ممکن است ناشی از ویژگی‌های آکوستیکی خیشومی‌شدگی واکه باشند. تمایزهای واکه‌ای در بافت‌های خیشومی، گرایش به حذف یا خنثی‌شدن دارند. به عنوان مثال، در بعضی از گویش‌های انگلیسی آمریکایی، pen و pin به تدریج یکسان شده‌اند. رایت^۱ (۱۹۸۶) (در گزارش خود اظهار می‌دارد که تأثیر خیشومی‌شدگی واکه، کوچک‌شدن کلی فضای واکه‌ای درکی است. در فضای درکی، واکه‌های افراشته خیشومی شده پایین‌تر از واکه‌های افراشته غیرخیشومی قرار دارند و واکه‌های افتاده خیشومی شده بالاتر از واکه‌های افتاده غیرخیشومی قرار دارند. این مشاهدات منجر به حدس و گمان‌هایی درباره منشأ آکوستیکی الگوهای تغییر واکه‌ای^۲ شد. واکه‌ها در بسیاری از زبان‌ها، شامل زبان انگلیسی، دستخوش تغییر زنجیره‌ای^۳ می‌شوند، بدین صورت که واکه‌ها یک گام افراشته‌تر می‌شوند و واکه‌های افراشته به واکه‌های مرکب تبدیل می‌شوند. خیشومی‌شدگی واکه‌ای ناتمايزدهنده ممکن است در به جریان‌انداختن این تغییرات

1. Wright

2. vowel shift

3. chain shift

زنجیره‌ای نقش داشته باشد، زیرا واکه‌های افتاده مستعد مقداری خیشومی‌شدگی بی‌اختیار^۱ هستند، بدین صورت که وقتی زبان برای تولید واکه افتاده پایین می‌آید، پرده نرم کام به وسیله عضله زبانی کامی^۲ کشیده و باز می‌شود (مول^۳، ۱۹۶۲؛ به لوب کر^۴، ۱۹۶۸ نیز رجوع کنید). این خیشومی‌شدگی بی‌اختیار ممکن است منجر به ارزیابی مجدد کیفیت واکه‌های افتاده به لحاظ درکی شود، زیرا با خیشومی‌شدگی واکه، سازه و پادسازه مجرای خیشوم نیز به آن اضافه می‌شود. از این رو، به دلیل تأثیرات آکوستیکی و درکی خیشومی‌شدگی واکه، یک تغییر زنجیره‌ای ممکن است به شکل یک زنجیره فشار^۵ آغاز شود.

۹-۵. درک همخوان خیشومی

شکل ۹-۱۱ نقشه درکی سه زبان را که توسط هارنس برگر^۶ (۲۰۰۱) بررسی شده‌اند، نشان می‌دهد. در این مطالعه، محرک‌ها، همخوان‌های خیشومی مشدیدی هستند که در بافت [i_i] توسط یک گویشور مالایالام^۷ تولید شده‌اند. شنوندگان، محرک‌ها را در دسته‌های سه‌تایی مانند زیر می‌شنیدند:

A	X	B
immi	injɪ	injɪ

از شنوندگان خواسته شد که بگویند آیا محرک میانی (X) بیشتر به محرک اولی (A) شباهت دارد یا به آخری (B). در مورد مثال بالا، شنوندگان اغلب (B) را انتخاب می‌کردند، زیرا خیشومی کامی [ɲ] شباهت بیشتری به خیشومی نرم کامی [ŋ] دارد تا به خیشومی دولبی [m]. می‌توانیم تعداد دفعاتی که شنوندگان [ɲ] و [ŋ] را به یکدیگر شبیه‌تر دانسته‌اند، به عنوان معیاری برای میزان نزدیکی آنها در فضای درکی در نظر بگیریم و از این داده‌ها برای رسم نقشه‌های درکی استفاده کنیم.

1. passive nasalization

2. palatoglossus

3. Moll

4. Lubker

5. push chain

6. Harnsberger

7. Malayalam. یکی از زبان‌های خانواده زبانی درآویدیان در کشور هندوستان. (مترجم)

نقشه‌های شکل ۹-۱۱ توسط «هارنسبرگر» رسم شده‌اند. من آنها را چرخاندم به‌طوری که بر روی محور عمودی مقدار [m] و [ŋ] یکسان شد. به نظر من، با این کار مقایسه فضاها درکی زبان‌های مختلف آسان‌تر می‌شود. این نقشه‌ها (که شامل تنها سه گروه از هفت گروهی که «هارنسبرگر» آزمایش کرد، هستند!) از نظر اطلاعاتی غنی هستند. ما در این نتایج فقط به تعدادی از جنبه‌های فضای درکی خیشومی‌ها می‌پردازیم. فراموش نکنید که برای رسم نقشه فضاها درکی، از محرک‌های کاملاً یکسانی استفاده می‌کنیم و تنها تفاوت یک نقشه نسبت به نقشه دیگر این است که شنوندگان به زبان‌های متفاوتی صحبت می‌کنند.

برخی از جنبه‌های کلی فضاها درکی در مالایالام، ماراتی^۱ و انگلیسی یکسان است (و برای سایر گروه‌های شنوندگان در مطالعه «هارنسبرگر» نیز یکسان بودند). به‌عنوان مثال، خیشومی‌های تیغه‌ای در نیمه بالایی فضاها درکی و خیشومی‌های لبی و بدنه‌ای در نیمه پایینی آن قرار دارند (البته موقعیت خیشومی‌های کامی روی این بُعد تا حدی مبهم است). محور افقی تا اندازه‌ای با محل تولید مطابقت دارد. برای تمام شنوندگان صرف نظر از نوع زبان بومی‌شان، [n] و [ŋ] به هم نزدیک هستند، همان‌طور که [n] و [ɳ] نیز به هم نزدیک‌اند.

این داده‌ها برخی از الگوهای جالب زبان‌ویژه در مورد درک خیشومی‌ها را نشان می‌دهند. اولاً به موقعیت همخوان برگشتی [n] در ماراتی توجه کنید که متفاوت از انگلیسی و مالایالام است. در انگلیسی و مالایالام، [n] بین خیشومی‌های دندانی و لثوی [n] و [ɳ] و خیشومی لبی [m] قرار دارد. در ماراتی، [n] در فاصله دورتری از [m] در مقایسه با [n] و [ɳ] قرار دارد. این فقط تصور من است اما ممکن است به این دلیل باشد که گویشوران انگلیسی و مالایالام به گذر افتان سازه سوم هجای VC در [in] در محرک [inɳi] توجه می‌کنند و بین این گذر و گذرهای سازه‌ای افتان هجای VC در [im] شباهت پیدا می‌کنند، در حالی که شنوندگان ماراتی به گذرهای رهش هجای CV در محرک [inɳi] توجه می‌کنند که در این صورت، بیشتر به خیشومی‌های لثوی و دندانی شباهت دارد.

۱. Marathi. یکی از زبان‌های شاخه هندوایرانی که در هند به آن صحبت می‌شود. (مترجم)

nn	ن ملايا لام
n η	m
n n	انگلیسی
n η	n
η	m
n n n	ماراتی
n η	m

شکل ۹-۱۱. نقشه‌های درکی خیشومی‌ها برای گویشوران زبان‌های مالایالام، انگلیسی و ماراتی. سایه خاکستری نشان‌دهنده این است که آوای مورد نظر در فهرست خیشومی‌های آن زبان موجود نیست. در آی‌پی‌ای، علائم [m, n, ɳ, ɲ, ɳ̌, ɳ̍] نشان‌دهنده خیشومی‌های دولبی، دندانی، لثوی، برگشتی، کامی و نرم کامی است. (ردراون^۱ اقتباس از هارنس‌برگر، ۲۰۰۱)

جنبه دیگری که زبان‌ویژه و بسیار قابل توجه در این شکل‌هاست، این است که دو حیثی‌نوی [n] و [ɲ] در فضای درکی مالایالام در هم ادغام شده‌اند. آشور و کوماری^۲ (۱۹۹۷) اظهار می‌دارند که اگرچه [n] و [ɲ] هر دو در مالایالام وجود دارند، اما تقریباً در توزیع تکمیلی هستند و «بار نقشی تقابل آنها کم است». بدیهی است که بار نقشی پایین تقابل [n] و [ɲ] در مالایالام سبب شده است که این دو آوا برای شنوندگان مالایالام شباهت بیشتری به یکدیگر داشته باشند در مقایسه با شنوندگانی که زبان‌شان تمایز بین این دو آوای مشابه به‌لحاظ آکوستیکی را تحمل می‌کند.

منابعی برای مطالعه بیشتر

- Asher, R. E. and Kumari, T. C. (1997) *Malayalam*, London: Routledge.
- منبعی برای مطالعه مقدماتی مالایالام برای تفسیر داده‌های در کی که در بخش ۹-۵ مطرح شد.
- Diehrn, B. and Johnson, K. (1997) Near-merger in Russian palatalization. *OSU Working Papers in Linguistics*, 50, 1 1-18.
- نمونه‌ای از پیامدهای در کی فرایند ادغام جزئی^۱ که باز هم در رابطه با مثال مالایالام است.
- Fant, C. (1960) *Acoustic Theory of Speech Production*, The Hague: Mouton.
- تحلیل بنیادین و نوآورانه همخوان‌های خیشومی و کناری.
- Fujimura, O. (1962) Analysis of nasal consonants. *Journal of the Acoustical Society of America*, 32, 1865-75.
- مقاله‌ای کوتاه با ارائه تحلیلی آکوستیکی که از بوته آزمایش سربلند بیرون آمد.
- Harnsberger, J. D. (2001) The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics*, 29, 303-27.
- این مقاله به طرز استثنایی به معرفی درک چالش برانگیز و معنادار گفتار زبان گذر می‌پردازد.
- Lindqvist-Gauffin, J. and Sundberg, J. (1976) Acoustic properties of the nasal tract. *Phonetica*, 33, 161-8.
- فرکانس پادسازه‌هایی را که به وسیله سینوس‌های بینی ایجاد می‌شوند، اندازه‌گیری کرد.
- Lubker, J. (1968) An EMG-cinefluorographic investigation of velar function during normal speech production. *Cleft Palate Journal*, 5, 1-1 8.
- داده‌هایی که نشان می‌دهد واکه‌های افتاده مقداری خیشومی شدگی خودبه‌خود دارند.
- Maeda, S. (1993) Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In M. K. Huffman and R. A. Krakow (eds.), *Phonetics and Phonology*, vol. 5: *Nasals, nasalization, and the velum*, New York: Academic Press, 147-67.
- یک مقاله نظری و تجربی مهم در زمینه درک ما از خیشومی شدگی واکه.
- Marple, L. (1987) *Digital Spectral Analysis with Applications*, Englewood Cliffs, NJ: Prentice Hall.
- حاوی بحث مفیدی راجع به مدل قطب-صفر^۲ طیف‌ها (از این رو، می‌توانید هم سازه‌ها و هم پادسازه‌ها را اندازه بگیرید).
- Moll, K. L. (1962) Velopharyngeal closure in vowels. *Journal of Speech and Hearing Research*, 5, 30-7.
- یک دیدگاه تولیدی راجع به خیشومی شدگی خودبه‌خود واکه افتاده.
- Petursson, M. (1973) Quelques remarques sur l'aspect articulatoire et acoustique des constrictives intrabuccales Islandaises. *Travaux de l'Institut de Phonétique de Strasbourg*, 5, 79-99.
- ردیابی اشعه X از همخوان‌های ايسلندی.
- Qi, Y. (1989) Acoustic features of nasal consonants. Unpublished Ph.D. diss., Ohio State University.
- نمونه‌ای از کاربرد مدل قطب-صفر برای تعیین فرکانس پادسازه‌ها در همخوان‌های خیشومی.

1. near merger

۲. pole-zero. قطب همان سازه و صفر همان پادسازه است. (مترجم)

Rayleigh, J. W. S. (1896) *The Theory of Sound*, London: Macmillan; repr. 1945, New York: Dover.

مقدمه‌ای بر آکوستیک با بحثی اولیه در مورد پادبازخوان‌های آکوستیکی.

Repp, B. (1986) Perception of the [m]-[n] distinction in CV syllables. *Journal of the Acoustical Society of America*, 79, 1987-99.

درک محل تولید در خیشومی‌ها، به‌ویژه نشان می‌دهد که گذرهای سازه‌ای اهمیت حیاتی دارند.

Stone, M. (1991) Toward a model of three-dimensional tongue movement. *Journal of Phonetics*, 19, 309-20.

استفاده از امواج فراصوت برای ردیابی حرکات از یک سمت به سمت دیگر زبان در همخوان‌های کناری.

Straka, G. (1965) *Album phonétique*, Laval: Les Presses de l'Université Laval.

مجموعه مفیدی از ردیابی‌های اشعه X که تولیدهای متعدد گفتار را نشان می‌دهد.

Wright, J. T. (1986) The behavior of nasalized vowels in the perceptual vowel space. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental Phonology*, New York: Academic Press, 45-67.

یکی از منابع کلیدی درخصوص منشا آوایی-درکی تغییرات ارتفاع واکه در واکه‌های خیشومی‌شده.

تمرین

واژگان تخصصی

این اصطلاحات را تعریف کنید: پهنای باند، میرایی، حفره یا شاخه جانبی در دستگاه گفتار، پادبازخوان، پادسازه، «حاشیه» بازخوان، سینوس، مجرای دهانی، مجرای خیشوم، دریچه کامی حلقی، خیشومی شدگی بی اختیار.

پرسش

۱. در شکل ۹-۲ حاشیه کدام یک شیب بیشتری دارد: شکل موجی که میرایی شدید دارد یا شکل موجی که میرایی خفیف دارد؟
۲. بازخوان‌های مجرای خیشوم را اگر طول حفره حلق ۸ سانتی‌متر و طول حفره خیشوم ۱۲ سانتی‌متر باشد، به دست آورید.
۳. مقدار فرکانس A_1 را در یک انسدادی خیشومی کامی به دست آورید.
۴. در مورد ویژگی‌های آکوستیکی همخوان‌های کناری کامی و نرم‌کامی با توجه به مبحث این فصل در زمینه همخوان‌های کناری لثوی، نظر دهید. آیا انتظار دارید این آواها پادسازه داشته باشند؟
۵. توضیح دهید که چگونه واکه‌هایی که در آنها چاکنای بازتر از حد معمول است (واکه‌های واک‌دار نفسی یا بخشی از یک واکه که مجاور آوایی با جریان هوای زیاد، مثلاً یک سایشی بی‌واک قرار دارد) ممکن است در معرض «خیشومی شدگی خودبه‌خود»^۱ باشد. مثالی از خیشومی شدگی خودبه‌خود: کلمه فرانسوی *rosse* (با پسوند جمع -ed) به معنی «اسب‌ها» به صورت *roncet* وارد زبان برتانی میانه^۲ شد. یک نکته: توجه کنید که چگونه تأثیر آکوستیکی افزایش بیش از حد معمول مرحله باز چاکنای هنگام ارتعاش، ممکن است شبیه تأثیر بازشدن بینی باشد.

1. spontaneous nasalization

۲. Middle Breton. زبانی در شمال کشور فرانسه که بین قرون ۱۱ تا ۱۶ میلادی به آن صحبت می‌شده است. (مترجم)

منابع

- Asher, R. E. and Kumari, T. C. (1997) *Malayalam*, London: Routledge.
- Best, C. T. (1995) A direct realist perspective on cross-language speech perception. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research*, Timonium, MD: York Press, 167—200.
- Bladon, A. and Lindblom, B. (1981) Modeling the judgment of vowel quality differences. *Journal of the Acoustical Society of America*, 69, 14 14—22.
- Bless, D. M. and Abbs, J. H. (1983) *Vocal Fold Physiology: Contemporary research and clinical issues*, San Diego: College Hill Press.
- Bond, Z. S. (1999) *Slips of the Ear: Errors in the perception of casual conversation*, San Diego: Academic Press.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, MA: MIT Press.
- Brödel, M. (1946) *Three Unpublished Drawings of the Anatomy of the Human Ear*, Philadelphia: Saunders.
- Campbell, R. (1994) Audiovisual speech: Where, what, when, how? *Current Psychology of Cognition*, 13, 76—80.
- Carford, J. C. (1977) *Fundamental problems in Phonetics*, Bloomington: Indiana University Press.
- Chiba, T. and Kajiyama, M. (1941) *The Vowel: Its nature and structure*, Tokyo: Kaiseikan.
- Cole, R. A. (1973) Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics*, 13, 153—6.
- Cooley, J. W., Lewis, P. A. W., and Welch, P. D. (1969) The fast Fourier transform and its applications. *IEEE Transactions on Education*, 12, 2 7—34.
- Cooper, F. S., Liberman, A. M., and Borst, J. M. (1951) The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of science*, 37, 318—25.
- Davis, S. and Mermelstein, P. (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP 28, 357—66.

- Delattre, P. C., Libermann, A. M., and Cooper, F. S. (1955) Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 27, 769—73.
- Egan, J. p. and Hake, H. W. (1950) On the masking pattern of a simple auditory stimulus. *Journal of the Acoustical Society of America*, 22, 622—30.
- Elman, J. L. and McClelland, J. L. (1988) Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27, 143—65.
- Fant, C. (1960) *Acoustic Theory of Speech Production*, The Hague: Mouton.
- Flanagan, J. L. (1965) *Speech Analysis Synthesis and Perception*, Berlin: Springer-Verlag.
- Flege, J. E. (1995) Second language speech learning: Theory, findings, and problems. In W Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research*, Timonium, MD: York Press, 167—200.
- Forrest, K., Weismer, &, Milenkovic, P., and Dougall, R. N. (1988) Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84, 115—23.
- Fry, D. B. (1979) *The Physics of Speech*, Cambridge: Cambridge University Press.
- Fujimura, O. (1962) Analysis of nasal consonants. *Journal of the Acoistical Society of America*, 32, 1865—75.
- Ganong, W. F. (1980) Phonetic categorization in auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 110—25.
- Green, K. P., Kuhl, P. K., Meltzoff, A. N., and Stevens, E. B. (1991) Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, 50, 524—36.
- Guion, S. C. (1998) The role of perception in the sound change of velar palatalization. *Phonetica*, 55, 18—52.
- Hagiwara, R. (1995) Acoustic realizations of American /r/ as produced by women and men. *UCLA Working Papers in Phonetics*, 90, 1—187.
- Halle, M. and Stevens, K. N. (1969) On the feature ‘advanced tongue root.’ *Quarterly Progress Report*, 94, 209—15. Research Laboratory of Electronics, MIT.
- Harnsberger, J. D. (2001) The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics*, 29, 303—27.
- Heinz, J. M. and Stevens, K. N. (1961) On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America*, 33, 589—96.
- Jakobson, R., Fant, C., and Halle, M. (1952) *Preliminaries to Speech Analysis*, Cambridge, MA: MIT Press.

- jassem, W. (1979) Classification of fricative spectra using statistical discriminant functions. In B. Lindblom and S. Öhman (eds.) *Frontiers of Speech Communication Research*, New York: Academic Press, 77—91.
- Johnson, K. (1989) Contrast and normalization in vowel perception. *Journal of Phonetics*, 18, 229—54.
- Johnson, K. (1992) Acoustic and auditory analysis of Xhosa clicks and pulmonic. *UCLA Working Papers in Phonetics*, 83, 33—47.
- Johnson, K. (2008) *Quantitative Methods in Linguistics*, Oxford: Wiley-Blackwell.
- Johnson, K. and Ralston, J. V. (1994) Automaticity in speech perception: Some speech/nonspeech comparisons. *Phonetica*, 51(4), 195—209.
- Johnson, K., Ladefoged, P., and Lindau, M. (1993) Individual differences in vowel production *Journal of the Acoustical Society of America*, 94, 701—14.
- Joos, M. (1948) Acoustic phonetics. *Language*, 23, suppl. 1.
- Klatt, D. H. and Klatt, L. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820-57.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992) Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science*, 255, 606-8.
- Ladefoged, P. (1996) *Elements of Acoustic Phonetics*, 2nd edn., Chicago: University of Chicago Press.
- Ladefoged, P. and Maddieson, I. (1996) *The Sounds of the World's Languages*, Oxford: Blackwell.
- Ladefoged, P., DeClerk, J., Lindau, M., and Papcun, G. (1972) An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics*, 22, 48—75.
- Lambacher, S., Martens, W., Nelson, B., and Berman, J. (2001) Identification of English voiceless fricatives by Japanese listeners: The influence of vowel context on sensitivity and response bias. *Acoustic Science & Technology*, 22, 334—43.
- Laver, J. (1980) *The Phonetic Description of Voice Quality*, Cambridge: Cambridge University Press.
- Liberman, A. M., Harris, K. S., Hoffman H. S., and Griffith, B. C. (1957) The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358—68.
- Liljencrants, J. and Lindblom, B. (1972) Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48, 83 9—62.
- Lindau, M. (1978) Vowel features, *Language*, 54, 541—63.
- Lindau, M. (1979) The feature “expanded.” *Journal of Phonetics*, 7, 163—76.
- Lindau, M. (1984) Phonetic differences in glottalic consonants. *Journal of Phonetics*, 12, 147-55.

- Lindau, M. (1985) The story of /r/. In V. Fromkin (ed.), *Phonetic Linguistics: Essays in honor of Peter Ladefoged*, Orlando, FL: Academic Press.
- Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H&H theory. in W. J. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modeling*, Dordrecht: Kluwer, 403—39.
- Lindqvist-Gauffin, J. and Sundberg, J. (1976) Acoustic properties of the nasal tract. *Phonetica*, 33, 161—8.
- Lotto, A.J. and Kluender, K. R. (1998) General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics*, 60, 602—19.
- Lubker, J. (1968) An EMG-cinefluorographic investigation of velar function during normal speech production. *Cleft Palate Journal*, 5, 1—18.
- Lyons, R. F. (1982) A computational model of filtering, detection and compression in the cochlea. *Proceedings of the IEEE international Conference on Acoustics, Speech and Signal Processing*, 1282—5.
- Lyons, R. F. (1997) *Understanding Digital Signal Processing*, Reading, MA: Addison-Wesley.
- Maddieson, I. (1984) *Patterns of Sounds*, Cambridge: Cambridge University Press.
- Maeda, S. (1993) Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In M. K Huffman and R. A. Krakow (eds.) *Phonetics and Phonology*, vol. 5: Nasals, nasalization, and the velum, New York: Academic Press, 147—67.
- Mann, V. A. (1980) Influence of preceding liquid on stop-consonant perception. *Perception & Psychophysics*, 28, 407—12.
- Marple, L. (1987) *Digital Spectral Analysis with Applications*, Englewood Cliffs, NJ: Prentice Hall.
- McGurk, H. and MacDonald, J. (1976) Hearing lips and seeing voices. *Nature*, 264, 746—8. McDonough, J. (1993) The phonological representation of laterals. *UCLA Working Papers in Phonetics*, 83, 19—32.
- McDonough, J. and Ladefoged, P. (1993) Navajo stops. *UCLA Working Papers in Phonetics*, 84, 151—64.
- Miller, G. A. and Nicely, P. E. (1955) An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338—52.
- Miller, J. D. (1989) Auditory perceptual interpretation of the vowel. *Journal of the Acoustical Society of America*, 85, 2114—34.
- Moll, K. L. (1962) Velopharyngeal closure in vowels. *Journal of Speech and Hearing Research*, 5, 30—7.
- Moore, B. C. J. (1982) *An Introduction to the Psychology of Hearing*, 2nd edn., New York: Academic Press.

-
- Moore, B. C. J. and Glasberg, B. R. (1983) Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America*, 74, 750—3.
- Mrayati, M., Carré, R., and Guérin, B. (1988) Distinctive regions and modes: A new theory of speech production. *Speech Communication*, 7, 257—86.
- O'Shaughnessy D. (1987) *Speech Communication: Human and machine*, Reading, MA: Addison-Wesley.
- Parzen, E. (1962) On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33, 1065—76.
- Pastore, R. E. and Farrington, S. M. (1996) Measuring the difference limen for identification of order of onset for complex auditory stimuli. *Perception & Psychophysics*, 58(4), 510—26.
- Patterson, R. D. (1976) Auditory filter shapes derived from noise stimuli. *Journal of the Acoustical Society of America*, 59, 640—54.
- Perkell, J. (1971) Physiology of speech production: A preliminary study of two suggested revisions of the features specifying vowels. *Quarterly Progress Report*, 102, 123—39. Research Institute of Electronics, MIT.
- Petursson, M. (1973) Quelques remarques sur l'aspect articulatoire et acoustique des constrictives intrabuccales Islandaises. *Travaux de l'Institut de Phonétique de Strasbourg*, 5, 79—99.
- Pickles, J. O. (1988) *An Introduction to the Physiology of Hearing*, 2nd edn., New York: Academic Press.
- Pisoni, D. B. (1977) Identification and discrimination of the relative onset time of two-component tones: Implications for voicing perception in stops. *Journal of the Acoustical Society of America*, 61, 1352—61.
- Potter, R. K., Kopp, C. A., and Green H. (1947) *Visible Speech*, Dordrecht: Van Nostrand.
- Qi, Y. (1989) Acoustic features of nasal consonants. Unpublished Ph.D. diss., Ohio State University.
- Rand, T. C. (1974) Dichotic release from masking for speech. *Journal of the Acoustical Society of America*, 55(3), 678—80.
- Raphael, L. J. and Bell-Berti, F. (1975) Tongue musculature and the feature of tension in English vowels. *Phonetica*, 32, 61—73.
- Rayleigh, J. W. S. (1896) *The Theory of Sound*, London: Macmillan; repr. 1945, New York: Dover.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., and Carrell, T. D. (1981) Speech perception without traditional speech cues. *Science*, 212, 947—50.
- Repp, B. (1986) Perception of the [m]-[n] distinction in CV syllables. *Journal of the Acoustical Society of America*, 79, 1987—99.
- Rosenblum, L. D., Schmuckler, M. A., and Johnson, J. A. (1997) The McGurk effect in infants. *Perception & Psychophysics*, 59, 347—57.

- Samuel, A. G. (1991) A further examination of the role of attention in the phonemic restoration illusion. *Quarterly journal of Experimental Psychology*, 43A, 679—99.
- Schroeder, M. R., Atal, B. S, and Hall, J. L. (1979) Objective measure of certain speech signal degradations based on masking properties of human auditory perception. In B. Lindblom and S. Ohman (eds.), *Frontiers of Speech Communication Research*, London: Academic Press, 217—29.
- Sekiyama, K. and Tohkura, Y. (1993) Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics*, 21, 427—44.
- Seneff, S. (1988) A joint synchrony/mean-rate model of auditory speech processing. *Journal of Phonetics*, 16, 55—76.
- Shadle, C. (1985) The acoustics of fricative consonants. *RLE Technical Report*, 506, MIT.
- Shadle, C. (1991) The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, 19, 409—24.
- Shannon, C. E. and Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana: University of Illinois.
- Shepard, R. N. (1972) Psychological representation of speech sounds. in E. E. David and P. B. Denes (eds.), *Human Communication: A unified view*. New York: McGraw-Hill, 67—113.
- Slaney, M. (1988) Lyons' cochlear model. *Apple Technical Report*, 13. Apple Corporate Library, Cupertino, CA.
- Stevens, K. N. (1972) The quantal nature of speech: Evidence from articulatory-acoustic data. In E. E. David, Jr. and P. B. Denes (eds.), *Human Communication: A unified view*, New York: McGraw-Hill, 51—66.
- Stevens, K. N. (1987) Interaction between acoustic sources and vocal-tract configurations for Consonants. *Proceedings of the Eleventh International Conference on Phonetic Sciences*, 3, 385—9.
- Stevens, K. N. (1989) On the quantal nature of speech. *Journal of Phonetics*, 17, 3—45.
- Stevens, K. N. (1999) *Acoustic Phonetics*, Cambridge, MA: MIT Press.
- Stevens, S. S. (1957) Concerning the form of the loudness function. *Journal of the Acoustical Society of America*, 29, 603—6.
- Stockwell, R. P. (1973) Problems in the interpretation of the Great English Vowel Shift. In M. E. Smith (ed.), *Studies in Linguistics in Honor of George L. Trager*, The Hague: Mouton, 344—62.
- Stone, M. (1991) Toward a model of three-dimensional tongue movement. *Journal of Phonetics*, 19, 309—20.
- Straka, G. (1965) *Album phonétique*, Laval: Les Presses de l'Université Laval.

-
- Syrdal, A. K. and Gophal, H. S. (1986) A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79, 1086—1100.
- Terbeek, D. (1977) A cross-language multidimensional scaling study of vowel perception. *UCLA Working Papers in Phonetics*, 37, 1—271.
- Traunmüller, H. (1981) Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America*, 69, 1465—75.
- Walker, S., Bruce, V., and O'Malley, C. (1995) Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect. *Perception & Psychophysics*, 57, 1124—33.
- Warren, R. M. (1970) Perceptual restoration of missing speech sounds. *Science*, 167, 392—3.
- Wright, J. T. (1986) The behavior of nasalized vowels in the perceptual vowel space. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental Phonology*, New York: Academic Press, 45-67.
- Zwicker, E. (1961) Subdivision of the audible frequency range into critical bands (*Frequenzgruppen*). *Journal of the Acoustical Society of America*, 33, 248.
- Zwicker, E. (1975) Scaling. In W. D. Keidel and W. D. Neff (eds.), *Auditory System: Physiology (CNS), behavioral studies, psychoacoustics*. Berlin: Springer-Verlag.

پاسخ به منتخبی از سؤالات

مبانی آکوستیک و صافی‌های آکوستیکی

فصل ۱

سؤال ۲

۱ ثانیه = ۱۰۰۰ میلی ثانیه

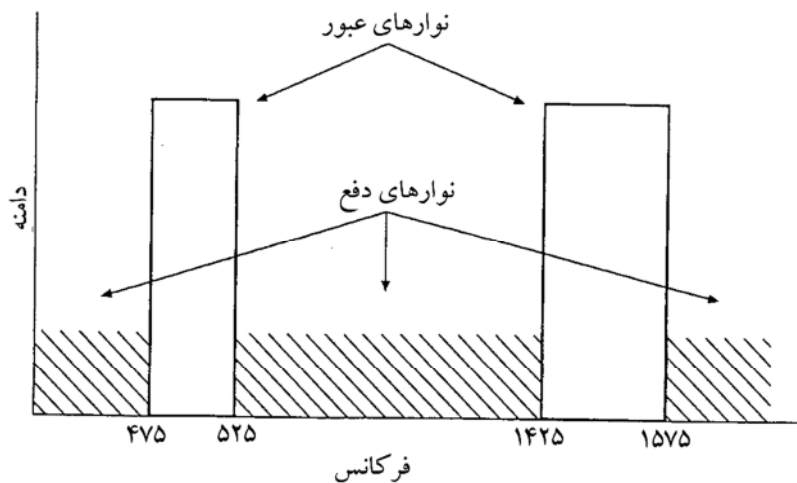
۰/۲ ثانیه = ۲۰۰ میلی ثانیه

۰/۰۱ ثانیه = ۱۰ میلی ثانیه

۱/۲۱ ثانیه = ۱۲۱۰ میلی ثانیه

سؤال ۷

نوار عبور و نوار دفع این صافی دوبل در شکل (الف) نشان داده شده است.



شکل الف ۱. نوارهای عبور و دفع صافی‌ای که در سؤال ۷ در فصل ۱ توصیف شد.

فصل ۲ نظریه آکوستیکی تولید گفتار

سؤال ۱

۳۸- دسی بل = دامنه در ۱ کیلوهرتز

$$\frac{۴۱\text{-دسی بل}}{۳\text{-دسی بل}} = \text{دامنه در } ۲ \text{ کیلوهرتز}$$

به نظر می رسد طیف شیب پیدا می کند، به طوری که افت دامنه در این بازه فرکانسی بر مبنای ۳ دسی بل است.

سؤال ۳

با استفاده از فرمول

جدول (الف ۱). فرکانس های سازه ای (برحسب هرتز) برای دستگاه گفتاری با طول های ۱۲، ۱۵ و ۱۸ سانتی متر. مقدار سازه ها با اصلاح طول $L = L + 1.2$ در پرانتز آمده است.

	n	(۲n-۱)c	L=۱۲	۱۵	۱۸
F ₁	1	35000	729 (663)	583 (540)	486 (456)
F ₂	2	105000	2188 (1989)	1750 (1620)	1458 (1367)
F ₃	3	175000	3646 (3314)	2917 (2701)	24331 (2279)

فصل ۳ پردازش دیجیتالی سیگنال

سؤال ۲

مقادیر فرکانس نایکویست به صورت زیر است:

۸۰۰۰ هرتز، ۵۵۰۱/۲۵ هرتز، ۱۰ هرتز

سؤال ۵

از آنجا که در روش خودهمبستگی، ایده آل ترین طول پسافت برای محاسبه زیرویمی، مقدار واقعی دوره تناوب سیگنال است (همان کاری که برای محاسبه زیرویمی انجام می دهیم)، بنابراین بهترین دیرش پسافت برابر با $1/F_0$ است.

$$\text{ثانیه } ۰/۰۰۴۹ = ۱/۲۰۴, \text{ ثانیه } ۰/۰۰۵ = ۱/۲۰۰, \text{ ثانیه } ۰/۰۱ = ۱/۱۰۰$$

سؤال ۸

طول پنجره مساوی با: $(0/02327) = 512/22000$ ثانیه است. در واقع منظور از اصطلاح «۲۲» کیلو یا «۲۲ کیلوهرتز» تقریباً همیشه نرخ نمونه برداری ۲۲۰۵۰ هرتز است که اگر آن را بکار ببریم، اندکی پاسخ‌های بدست آمده تغییر می‌کند.

سؤال ۹

۲۲۰۰۰ نمونه در ثانیه، ضربدر ۰/۰۲ ثانیه، برابر با ۴۴۰ نمونه است.

سؤال ۱۰

فاصله بین نقاط در طیف افافتی برابر با $21/53$ هرتز است. این مقدار با تقسیم فرکانس نایکویست که نصف نرخ نمونه برداری است (فرض کنیم ۱۱۰۲۵ هرتز) بر ۲۵۶ به دست می‌آید (یعنی $1/2$ پنجره افافتی و نیم دیگر آن بخش «موهومی»^۱ طیف را کدبندی می‌کند).

فصل ۴ مبانی شنوایی

سؤال ۱

شش سون = ۵۰۰۰۰ میکروپاسکال. بلندی ذهنی دوبرابر، مساوی با ۱۲ سون یعنی ۱۳۷۰۰۰ میکروپاسکال است. این صدا که بلندی ذهنی آن دوبرابر است، سطح فشار صوتی دارد که تقریباً سه برابر بزرگ‌تر است.

سؤال ۳

$22 \text{ بارک} = 7 \text{ کیلوهرتز}$	$8/5 \text{ بارک} = 1 \text{ کیلوهرتز}$
$22/5 \text{ بارک} = 8 \text{ کیلوهرتز}$	$13 \text{ بارک} = 2 \text{ کیلوهرتز}$
$0/5 \text{ بارک}$	$4/5 \text{ بارک}$

1. imaginary part

فصل ۵ درک گفتار

سؤال ۴

با زیرماتریس‌های [d] و [ð]

	d	ð
d	۰/۷۲۷	۰
ð	۰/۰۱۵	۰/۵۱۵

می‌توان میزان شباهت (S) را محاسبه نمود،

$$S_{d\dot{o}} = (0/0 + 0/015) / (0/727 + 0/515) = 0/0124$$

و سپس فاصله درکی (d) را،

$$\begin{aligned} d_{d\dot{o}} &= \ln(S_{d\dot{o}}) \\ &= \ln(0/0124) \\ &= 4/39 \end{aligned}$$

که بزرگ‌تر از فاصله بین [z] و [ð] است.

$$\begin{aligned} d_{z\dot{o}} &= \ln(S_{z\dot{o}}) \\ &= \ln(0/081) \\ &= 2/51 \end{aligned}$$

سؤال ۵

داده‌هایی که ترکیب فضاهای درکی شنیداری و دیداری با هم را نشان می‌دهند و پیش‌بینی اثر شنیداری-دیداری مک‌گراگ را میسر می‌سازند، به وضوح با این دیدگاه مطابقت دارند که شنوندگان، بستارهای گفتاری را درک می‌کنند که به وسیله هم شنوایی و هم بینایی منتقل می‌شوند. واضح است که «مطابقت» مساوی با اثبات نیست، اما دانشمندان تمایل به پذیرفتن نظریه‌ای را دارند که با داده‌های بیشتری هماهنگی دارد.

فصل ۶ واکه‌ها

سؤال ۲

در فرمول (۶.۲)، با مقدار $A_b = ۳$ و $l_c = ۲$ مقدار $A_b l_c = ۲۴$ و مقدار $۵۵۷۰/۴$ $\pi/۲$ است. بنابراین، برای مقادیر A_c بین $۰/۰۵$ و ۲ داریم:

A_c	$۰/۰۵$	$۰/۱$	$۰/۱۵$	$۰/۲$
F_1	۲۵۴	۳۶۰	۴۴۰	۵۰۹

وقتی مقدار A_c برابر با صفر است، F_1 نیز مساوی صفر است.

سؤال ۶

در طیف واکه‌های افتاده و متوسط مزاتک، بین قله‌های F_0 (اولین همساز) و F_1 می‌توان یک قله برای دومین همساز نیز مشاهده نمود. برخلاف طیف‌های ال‌پی‌سی در شکل (۶.۱۰ الف) که در آن پایین‌ترین قله همواره F_1 است، در طیف‌های شنیداری شکل (۶.۱۰ ب) پایین‌ترین بازخوان دستگاه گفتار (F_1) ممکن است دومین یا سومین قله در طیف باشد. اگر زیروبمی واک بالا باشد، حتی ممکن است ملاحظه شود که F_0 و F_1 فقط یک قله در طیف شنیداری تشکیل داده‌اند. از این رو، به نظر می‌رسد که اگر F_1 می‌تواند اولین، دومین یا سومین قله در طیف باشد، طراحی یک سازه‌یاب خودکار برای یافتن F_1 در طیف‌های شنیداری دشوار باشد.

فصل ۷ سایشی‌ها

سؤال ۲

هنگامی که اندام‌های تولیدی در شروع یک سایشی به سمت هم می‌آیند و نیز هنگامی که گرفتگی در انتهای یک سایشی رهش پیدا می‌کند، احتمالاً این فرض که حفره‌های جلو و عقب جفت نمی‌شوند، نادرست است. زیرا گرفتگی، در این

لحظه‌ها، انسداد کمتری دارد. از این رو، انتظار می‌رود در آغاز و پایانه یک سایشی، قله‌های طیفی در فرکانس‌های بازخوانی حفره عقبی نیز تولید شوند.

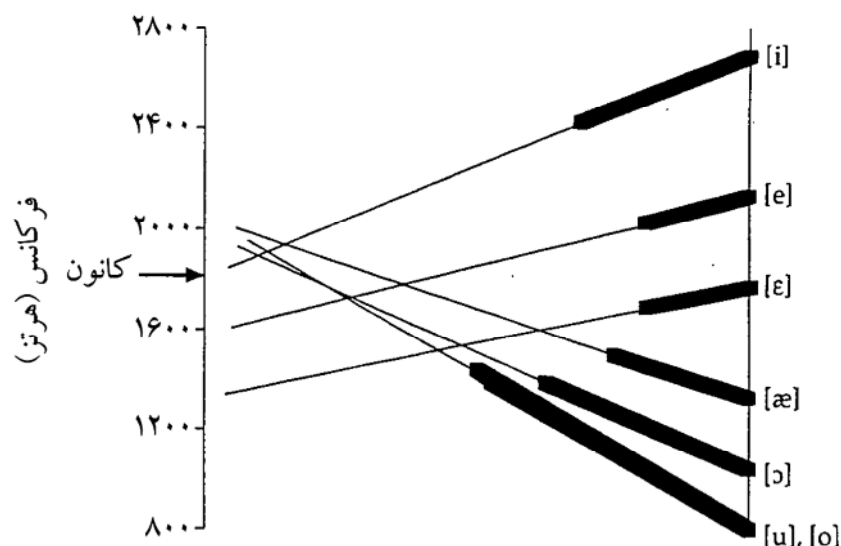
سؤال ۴

سؤال برای برخی دانشجویان عجیب است. کلید پاسخ، جفت‌شدگی آکوستیکی است. وقتی حفره‌های جلو و عقب کاملاً جفت می‌شوند (همان‌طور که در واکه‌ها می‌شوند)، بازخوان‌های آنها برای ایجاد نواحی پایدار ادغام می‌شود. اما وقتی حفره‌های جلو و عقب کاملاً جفت نمی‌شوند (مثل سایشی‌ها)، بازخوان‌های حفره جلویی از روی بازخوان‌های حفره عقبی می‌پرند (شکل ۷.۶). بنابراین، اگر نواحی ذره‌ای، مشخصه‌های تمایزدهنده را تعیین می‌کنند و اگر همان‌طور که شکل‌ها نشان می‌دهند، نواحی ذره‌ای برای محل تولید سایشی‌ها و واکه‌ها متفاوت باشد، پس مشخصه‌های محل تولید واکه‌ها و سایشی‌ها باید متفاوت باشد. این نتایج تحت این نظر کلی قرار می‌گیرند که بر طبق ملاحظات آکوستیکی، واکه‌ها و همخوان‌ها مشخصه‌های متفاوتی دارند، در حالی که بر طبق ملاحظات تولیدی، می‌توان واکه‌ها و همخوان‌ها را با مشخصه‌های یکسانی توصیف نمود.

فصل ۸ انسدادی‌ها و انسایشی‌ها

سؤال ۲

شکل (الف ۲) تصویری را نشان می‌دهد که با اندازه‌گیری گذرهای سازه‌ای شکل ۸.۷ به دست آمده است. خط عمودی، نشان‌دهنده نقطه شروع وضعیت ایستادن واکه در هر یک از هجاهای dV است. خط‌های پررنگ، گذرهای سازه‌ای و خط‌های کم‌رنگ، امتداد گذرها به سمت عقب و تلاقی آنها با یکدیگر را نشان می‌دهند. من فرکانس کانونی را انتخاب کردم که وزن بیشتری به گذرهای بلندتر و راحت‌تر برای اندازه‌گیری [u] و [i] را می‌دهد.



شکل (الف ۲). فرکانس کانونی F_2 در همخوان [d].

سؤال ۴

بله، انسایشی‌ها انفجار رهش دارند. زیرا یک بخش انسدادی دارند، به این دلیل که همان افزایش فشار و رهشی که در انسدادی‌های غیرانسایشی می‌بینیم در انسایشی‌ها نیز وجود دارد. اما به هر حال، انفجار رهش انسایشی‌ها ممکن است به لحاظ شنیداری با شروع ناگهانی سایش که بلافاصله پس از رهش انسداد روی می‌دهد، مستتر شود.

فصل ۹ خیشومی‌ها و کناری‌ها

سؤال ۱

در شکل ۹-۲ قله‌ای که میرایی آن خفیف‌تر است، «حاشیه» تندتری دارد. انرژی در قله‌ای که میرایی آن شدید است، با سرعت کمتری نسبت به قله‌ای که میرایی آن خفیف است، افت پیدا می‌کند.

سؤال ۳

فرض کنید که حفره دهان در یک همخوان کامی خیشومی ۴ سانتی متر طول دارد. پس فرکانس پایین ترین پادسازه برابر است با:

$$\begin{aligned} A_1 &= \frac{c}{4L} \\ &= \frac{35000}{16} \\ &= 2188 \end{aligned}$$

سؤال ۵

آنچه در واکه‌های نفسی و واکه‌های خیشومی شده مشترک است، این است که در طیف هر دو، در مقایسه با طیف واکه‌های خیشومی نشده که با واک عادی تولید می‌شوند، انرژی بیشتری در فرکانس‌های پایین وجود دارد. طیف واک در واکه‌های نفسی شیب طیفی تندتری دارد و در نتیجه، انرژی نسبتاً بیشتری بر روی فرکانس‌های پایین مشاهده می‌شود. در واکه‌های خیشومی شده، دو پادسازه‌ها در واکه‌های وجود دارد (اولین سازه‌های «خیشومی» و «دهانی») و پادسازه‌ها در واکه‌های خیشومی شده، انرژی فرکانس‌های بالای طیف را تضعیف می‌کنند.

بنابراین، «خیشومی شدگی خودبه‌خود» ممکن است نتیجه یک ابهام درکی باشد، بدین صورت که شنوندگانی که واکه‌های نفسی را می‌شنوند، ممکن است به اشتباه تصور کنند که گوینده یک واکه خیشومی شده تولید کرده است و اگر واکه‌های خیشومی شده فقط در بافت یک همخوان خیشومی تولید می‌شوند، شنونده ممکن است واکه‌ای را که به اشتباه خیشومی شده تشخیص داده، به صورت توالی واکه – خیشومی، مشابه سایر واکه‌های خیشومی شده در زبان، تقطیع کند.

واژه‌نامه فارسی به انگلیسی

analog device	ابزار آنالوگ
digital device	ابزار دیجیتال
confusability	ابهام‌زایی
interpolation	اتصال
dissipation	اتلاف
perceptual magnet effect	اثر آهن‌ربای درکی
contrast effect	اثر تقابلی
Ganong effect	اثر گانونگ
McGurk effect	اثر مک‌گرک
near merger	ادغام جزئی
masking	استتار
backward masking	استتار پس‌رو
forward masking	استتار پیش‌رو
temporal masking	استتار زمانی
frequency masking	استتار فرکانسی
simultaneous masking	استتار هم‌زمان
oscilloscope	اسیلوسکوپ
saturation	اشباع
mispronunciation	اشتباه تلفظی
misperception	اشتباه درکی
misheard	اشتباه شنیداری
end correction	اصلاح طول
distorted	اعوجاج

turbulence	اغتشاش
obstacle turbulence	اغتشاش مانعی
channel turbulence	اغتشاش مجرای
periodicity	الگوی تناوبی
sawtooth pattern	الگوی دندان‌اره‌ای
MP3	ام‌پی‌تری
International Phonetic Association (IPA)	انجمن بین‌المللی آواشناسی (آی‌پی‌ای)
articulator	اندام تولیدی
potential energy	انرژی پتانسیل
kinetic energy	انرژی جنبشی
affricate	انسایشی
phonetic coherence	انسجام آوایی
stricture	انسداد
acoustic reflection	انعکاس آکوستیکی
release burst	انفجار رهش
steady-state	ایستا
product array	آرایه حاصل ضرب
absolute threshold	آستانه مطلق
disturbance	آشفستگی
syllable onset	آغاز هجا
onset	آغاز
categorical magnet	آهن‌ربای مقوله‌ای
aerodynamic	آیرودینامیک
bass	بَم
lossy	بافت
cross-mode resonance	بازخوان عرضی
Helmholtz resonator	بازخوان هلمهولتز
creaky	بازداشته
phoneme restoration	بازگردانی واج

imaginary part	بخش موهومی
aliasing	بدنمایی
spectral prominence	برجستگی طیفی
online	برخط
signal clipping	برش سیگنال
tongue retroflexion	برگشت زبان
retroflex	برگشتی
egressive	برونسو
greatest common denominator	بزرگ‌ترین مقسوم‌علیه مشترک
magnitude	بزرگی
gesture	بستار
posture	بستار
gesturalist	بستارگرایی
perceived loudness	بلندی درکی
subjective loudness	بلندی ذهنی
relative loudness	بلندی نسبی
loss-less	بی‌افت
burstless	بی‌انفجار
bit	بیت
anti-resonance	پادبازخوان
anti-formant	پادسازه
antinode	پادگره
anti-matter	پادماده
auditory response	پاسخ فَرَکانسی
syllable coda	پایانهٔ هجا
offset	پایانه
larynx lowering	پایین‌رفتنی حنجره
diffuse	پراکنده
digital signal processing (DSP)	پردازش دیجیتالی سیگنال (دی‌اس‌پی)

basilar membrane	پردهٔ بازیلار
lag	پس‌افت
analysis window	پنجرهٔ تجزیه
waveform window	پنجرهٔ شکل‌موج
window	پنجره
windowing	پنجره‌گذاری
envelope	پوش
dynamic	پویا
bandwidth	پهنای باند
equivalent rectangular bandwidth	پهنای باند مستطیلی هم‌ارز
lip protrusion	پیش‌آمدگی لب
pre-emphasis	پیش‌تقویت
first-difference pre-emphasis	پیش‌تقویت اولین اختلاف
pre-glottalized	پیش‌چاکنایی‌شده
pre-nasalized	پیش‌خیشومی‌شده
pre-aspirated	پیش‌دمیده
pixel	پیکسل
stimulus continuum	پیوستار محرکی
response function	تابع پاسخ
discrimination function	تابع تشخیص
identification function	تابع شناسایی
fast Fourier transform (FFT)	تبدیل سریع فوریه (افاف‌تی)
inverse discrete Fourier transform	تبدیل معکوس گسستهٔ فوریه
auditory scene analysis	تجزیهٔ صحنهٔ شنیداری
Fourier analysis	تجزیهٔ فوریه
hierarchical cluster analysis	تحلیل خوشه‌ای سلسله‌مراتبی
autoregressive moving average analysis	تحلیل میانگین متحرک خودبازگشتی
compression	تراکم
resonate	تشدیدشدن، بازخوان‌شدن

warped	تغییر شکل
chain shift	تغییر زنجیره‌ای
polarity change	تغییر قطبیت
vowel shift	تغییر واکه‌ای
cognitive parsing	تقطیع شناختی
amplify	تقویت کردن
spit spike	تکانه بزاقی
impulse	تکانه، پالس
inter-speaker variability	تنوع بینافردی
perceptual illusion	توهم درکی
strident	تیز
movement	جاب‌جایی
compensation for coarticulation	جبران هم‌تولیدی
turbulent flow	جریان اغتشاشی
vortice	جریان حلقوی
laminal flow	جریان لایه‌ای
turbulent airflow	جریان هوای اغتشاشی
vocal fry	جز جز چاکنایی
uncoupled	جفت نشده
tongue-root advancement	جلو کشیدن ریشه زبان
tongue bunching	جمع شدن زبان
onset licensing	جواز آغاز
perceptual licensing	جواز درکی
glottalized	چاکنایی شده
skirt	حاشیه
rocking motion	حرکت موجی
harmonic motion	حرکت هماهنگ
side cavity	حفره جانبی
sublingual cavity	حفره زیرزبانی

cochlea	حلزون شنوایی
cochleagram	حلزون نگاشت
voice bar	خط واک
nasalization	خیشومی شدگی
passive nasalization	خیشومی شدگی بی اختیار
spontaneous nasalization	خیشومی شدگی خودبه خود
acoustic amplitude	دامنه آکوستیکی
perceptual amplitude	دامنه درکی
peak amplitude	دامنه قله
peak-to-peak amplitude	دامنه قله به قله
overall amplitude	دامنه کلی
amplitude	دامنه
duplex perception	درک دوطرفه
speech perception	درک گفتار
categorical perception	درک مقوله‌ای
ingressive	درون سو
valley	دره
velopharyngeal port	دریچه کامی حلقی
chart recorder	دستگاه ثبت موج
decibel sensation level (dB SL)	دسی بل سطح احساس (دی بی اس ال)
decibel sound pressure level (dB SPL)	دسی بل سطح فشار صوتی (دی بی اس پی ال)
aspiration	دمش
aspirated	دمیده
pitch-doubling	دوبرابر شدن زیروبمی
period	دوره تناوب
pitch period	دوره تناوب زیروبمی
diplophonia	دو صدایی
intraoral	دهانی
oral	دهانی

digit	دیجیت
duration	دیرش
compact disc	دیسک فشرده
alveolar ridge	دیواره لثه
x-ray tracing	ردیابی اشعه X
sonorant	رِسا
acoustic medium	رسانای آکوستیکی
rarefaction	رقت
linear predictive coding (LPC)	رمزگذاری پیش‌گویانه خطی (ال‌پی‌سی)
cepstral coding	رمزگذاری کپسترال
root mean square (RMS)	ریشه میانگین مربعات (آرام‌اس)
palatoglossus	زبانی‌کامی
rise time	زمان افزایش
voice onset time (VOT)	زمان آغاز واک (وی‌اُتی)
nasal murmur	زمزمه خیشومی
push chain	زنجیره فشار
formant	سازه
formant tracker	سازه‌یاب
static	ساکن
psycho-acoustics	سایکواکوستیک
tenseness	سختی
velocity	سرعت
particle velocity	سرعت ذره‌ای
speed of sound	سرعت صوت
cue	سرنخ
silence	سکوت
speech synthesizer	سنتزکننده گفتار، گفتارساز
bias	سوگیری
sone	سون

CD	سی دی
CD-ROM	سی دی رام
CD player	سی دی خوان
sensory input system	سیستم درون داد حسی
peripheral auditory system	سیستم شنوایی بیرونی
simple harmonic system	سیستم هماهنگ ساده
cycle	سیکل
continuous signal	سیگنال پیوسته
discrete signal	سیگنال گسسته
sinusoidal	سینوسی
side branch	شاخه جانبی
perceptual similarity	شباهت درکی
intensity	شدت
acoustic intensity	شدت آکوستیکی
boundary condition	شرط مرزی
pulmonic	ششی
wave shape	شکل موج
configuration	شکل بندی
acoustic waveform	شکل موج آکوستیکی
line busy	شلوغی خط
filter slope	شیب صافی
high-pass filter	صافی بالاگذر
anti-aliasing filter	صافی پادبدنمایی
band-pass filter	صافی نوارگذر
zero padding	صفرگذاری
sibilant	صفیری
open quotient (oq)	ضریب بازی (اُکیو)
timbre	طنین
wavelength	طول موج

auditory spectrum	طیف شنیداری
power spectrum	طیف قدرت
spectrograph	طیف‌نگار
tongue retraction	عقب‌کشیدن زبان
glide	غلت
nonlinear	غیرخطی
nonsinusoidal	غیرسینوسی
non-sibilant	غیرصفیری
phase	فاز
perceptual distance	فاصله درکی
ultrasound	فراصوت
resonant frequency	فرکانس بازخوانی
natural resonant frequency	فرکانس بازخوانی طبیعی
fundamental frequency	فرکانس پایه
cutoff frequency	فرکانس قطع
locus frequency	فرکانس کانونی
center frequency	فرکانس مرکزی
Nyquist frequency	فرکانس نایکویست
subglottal pressure	فشار زیر حنجره‌ای
compact	فشرده
data compression	فشرده‌سازی داده
acoustic vowel space	فضای واکه‌ای آکوستیکی
ejective	فورانی
dorsal ejective	فورانی بدنه‌ای
low-pass filter	فیلتر پایین‌گذر
inverse filtering	فیلتر معکوس
Shepard's Law	قانون شپارد
Cooper's rule	قانون کوپر
Fourier's theorem	قضیه فوریه

broad spectral peak	قله‌های پهن طیفی
palatogram	کام‌نگاشت
palatalization	کامی‌شدگی
spectral tilt	کجی طیفی
quantization	کمّی‌سازی
lateral	کناری
formant transition	گذر سازه
transient	گذرا
node	گره
upward spread of masking	گسترش رو به بالای استتار
gestalt	گشتالت
careful speech	گفتار محتاطانه و با دقت
jittery	لرزان
articulatory slop	لغزش تولیدی
slip of the ear	لغزش گوش
natural log	لگاریتم طبیعی
Ruben's tube	لوله روبن
modes of vibration	مُدهای ارتعاش
confusion matrix	ماتریس ابهام
dissimilarity matrix	ماتریس عدم شباهت
time-varying	متغیربازمان
periodic	متناوب
triangulation	مثلث‌بندی
nasal tract	مجرای خیشومی
oral tract	مجرای دهانی
lateral channel	مجرای کناری
auto-correlation pitch tracking	محاسبهٔ زیروبمی با روش خودهمبستگی
landmark	محدوده
stimulus	محرك

base stimulus	محرک پایه
spectral smearing	محو طیفی
frequency smearing	محو فرکانسی
gradient	مدرج
cochlear model	مدل حلزونی
two-tube model	مدل دولوله‌ای
functional model	مدل کاربردی
tube model	مدل لوله‌ای
modulating	مدوله کردن
shutting stage	مرحله آمادگی
center of gravity	مرکز ثقل
area	مساحت
resting area	مساحت استراحتی
surface area	مساحت سطح
rectangular	مستطیلی
flat	مسطح
voicing quanta	مقادیر ذره ای واک
acoustic impedance	مقاومت آکوستیکی
aerodynamic impedance	مقاومت آیرودینامیکی
glottal impedance	مقاومت چاکنایی
acquired categoriality	مقوله‌بندی اکتسابی
Bark scale	مقیاس بارک
Mel scale	مقیاس مل
multidimensional scaling (MDS)	مقیاس‌گذاری چندبعدی (ام‌دی‌اس)
automatic gain control mechanism	مکانیسم کنترل خودکار بهره
McGurking	مک‌گرگ کردن
implosives	مکیده
wall source	منبع دیوار
lip source	منبع لب

obstacle source	منبع مانع
equal loudness contour	منحنی بلندی همسان
trajectory	منحنی گذر
incident wave	موج انتشار
reflected wave	موج انعکاسی
standing wave	موج ایستاده
velocity wave	موج سرعت
sine wave	موج سینوسی
sound wave	موج صوتی
longitudinal wave	موج طولی
transverse wave	موج عرضی
complex wave	موج مرکب
direct wave	موج مستقیم
alias component	مؤلفه بدنما
component	مؤلفه
damped	میرا، میراشدن
micro-pascal (μPa)	میکروپاسکال
noncontinuant	ناپیوسته
quantal region	ناحیه ذره‌ای
unaspirated	نادمیده
lateral approximant	ناسوده کناری
nonword	ناکلمه
aperiodic	نامتناوب
undamped	نامیرا
heterorganic	ناهم‌اندام
nonconsonantal	ناهمخوانی
whisper	نجوا
click	نچ‌آوا
sampling rate	نرخ نمونه‌برداری

signal to noise ratio	نسبت سیگنال به نوفه
pitch-halving	نصف‌شدن زیروبمی
perturbation theory	نظریه اختلال
adaptive dispersion theory	نظریه پراکندگی تطبیقی
quantal theory	نظریه ذره‌ای
source-filter theory of speech production	نظریه منبع-صافی تولید گفتار
breathy	نفسی
perceptual map	نقشه درکی
mapping	نگاشت
exponential	نمایی
amplitude trace	نمودار دامنه
pitch trace	نمودار زیروبمی
prototype	نمونه متعارف
sampling	نمونه‌برداری
regions of stability	نواحی پایدار
tone	نواخت
Pure tone	نواخت خالص
tone glide	نواخت غلتان
narrow-band	نوارباریک
reject band	نوار دفع
Pass-band	نوار گذر
pass band	نوار گذر
wide-band	نوارپهن
pressure fluctuation	نوسان فشار
phonation type	نوع واک‌سازی
chirp noise	نوفه چپ‌چپ‌های
quantization noise	نوفه کمی‌سازی
nomogram	نوموگرام
voicing	واک

creaky voicing	واک بازداشته
breathy voicing	واک نفسی
nasalized vowel	واکه خیشومی شده
monophthong	واکه ساده
non-nasalized vowel	واکه غیرخیشومی
preferred vowel	واکه مرجع
plain voiced vowel	واکه واکدار ساده
modal voice	واک عادی
vocalic	واکه‌ای
amplitude resolution	وضوح دامنه
time resolution	وضوح زمانی
frequency resolution	وضوح فرکانسی
task	وظیفه
discrimination task	وظیفه تشخیص
identification task	وظیفه شناسایی
vocalis	وکالیس
target	هدف
Hertz	هرتز
vowel harmony	هماهنگی واکه‌ای
acoustic correlate	همبسته آکوستیکی
harmonic	همساز
equalizer	همسان‌ساز
data smoothing	هموارسازی داده
moving average smoothing	هموارسازی میانگین متحرک
hamming	همینگ

واژه‌نامه انگلیسی به فارسی

A

absolute threshold	آستانهٔ مطلق
acoustic amplitude	دامنهٔ آکوستیکی
acoustic correlate	همبستهٔ آکوستیکی
acoustic impedance	مقاومت آکوستیکی
acoustic intensity	شدت آکوستیکی
acoustic medium	رسانای آکوستیکی
acoustic reflection	انعکاس آکوستیکی
acoustic vowel space	فضای واکه‌ای آکوستیکی
acoustic waveform	شکل موج آکوستیکی
acquired categoriality	مقوله‌بندی اکتسابی
adaptive dispersion theory	نظریهٔ پراکندگی تطبیقی
aerodynamic	آیرودینامیک
aerodynamic impedance	مقاومت آیرودینامیکی
affricate	انسایشی
alias component	مولفهٔ بدنما
aliasing	بدنمایی
alveolar ridge	دیوارهٔ لثه
amplify	تقویت کردن
amplitude	دامنه
amplitude resolution	وضوح دامنه
amplitude trace	نمودار دامنه
analog device	ابزار آنالوگ

analysis window	پنجره تجزیه
anti-aliasing filter	صافی پادبدنمایی
anti-formant	پادسازه
anti-matter	پادماده
antinode	پادگره
anti-resonance	پادبازخوان
aperiodic	نامتناوب
area	مساحت
articulator	اندام تولیدی
articulatory slop	لغزش تولیدی
aspirated	دمیده
aspiration	دمش
auditory response	پاسخ فرکانسی
auditory scene analysis	تجزیه صحنه شنیداری
auditory spectrum	طیف شنیداری
auto-correlation pitch tracking	محاسبه زیرویمی با روش خودهمبستگی
automatic gain control mechanism	مکانیسم کنترل خودکار بهره
autoregressive moving average analysis	تحلیل میانگین متحرک خودبازگشتی
B	
backward masking	استتار پسرو
band-pass filter	صافی نوارگذر
bandwidth	پهنای باند
Bark scale	مقیاس بارک
base stimulus	محرک پایه
basilar membrane	پرده بازیلار
bass	بَم
bias	سوگیری
bit	بیت
boundary condition	شرط مرزی

breathy	نفسی
breathy voicing	واک نفسی
broad spectral peak	قله‌های پهن طیفی
burstless	بی‌انفجار
C	
careful speech	گفتار محتاطانه و با دقت
categorical magnet	آهن‌ربای مقوله‌ای
categorical perception	درک مقوله‌ای
CD	سی‌دی
CD player	سی‌دی‌خوان
CD-ROM	سی‌دی‌رام
center frequency	فرکانس مرکزی
center of gravity	مرکز ثقل
cepstral coding	رمزگذاری کپسترال
chain shift	تغییر زنجیره‌ای
channel turbulence	اغتشاش مجرای
chart recorder	دستگاه ثبت موج
chirp noise	نوفه چپچه‌ای
click	نچ‌آوا
cochlea	حلزون شنوایی
cochleagram	حلزون‌نگاشت
cochlear model	مدل حلزونی
cognitive parsing	تقطیع شناختی
compact	فشرده
compact disc	دیسک فشرده
compensation for coarticulation	جبران هم‌تولیدی
complex wave	موج مرکب
component	مؤلفه
compression	تراکم

configuration	شکل‌بندی
confusability	ابهام‌زایی
confusion matrix	ماتریس ابهام
continuous signal	سیگنال پیوسته
contrast effect	اثر تقابلی
Cooper's rule	قانون کوپر
creaky	بازداشته
creaky voicing	واک بازداشته
cross-mode resonance	بازخوان عرضی
cue	سرنخ
cutoff frequency	فرکانس قطع
cycle	سیکل
D	
damped	میرا، میراشدن
data compression	فشرده‌سازی داده
data smoothing	هموارسازی داده
decibel sensation level (dB SL)	دسی‌بل سطح احساس (دی‌بی اس‌ال)
decibel sound pressure level (dB SPL)	دسی‌بل سطح فشار صوتی (دی‌بی اس‌پی‌ال)
diffuse	پراکنده
digit	دیجیت
digital device	ابزار دیجیتال
digital signal processing (DSP)	پردازش دیجیتالی سیگنال (دی‌اس‌پی)
diplophonia	دو صدایی
direct wave	موج مستقیم
discrete signal	سیگنال گسسته
discrimination function	تابع تشخیص
discrimination task	وظیفه تشخیص
dissimilarity matrix	ماتریس عدم شباهت
dissipation	اتلاف

distorted	اعوجاج
disturbance	آشفتگی
dorsal ejective	فورانی بدنه‌ای
duplex perception	درک دوطرفه
duration	دیرش
dynamic	پویا
E	
egressive	برون‌سو
ejective	فورانی
end correction	اصلاح طول
envelope	پوش
equal loudness contour	منحنی بلندی همسان
equalizer	همسان‌ساز
equivalent rectangular bandwidth	پهنای باند مستطیلی هم‌ارز
exponential	نمایی
F	
fast Fourier transform (FFT)	تبدیل سریع فوریه (افاف‌تی)
filter slope	شیب صافی
first-difference pre-emphasis	پیش‌تقویت اولین اختلاف
flat	مسطح
formant	سازه
formant tracker	سازه‌یاب
formant transition	گذر سازه
forward masking	استتار پیشرو
Fourier analysis	تجزیه فوریه
Fourier's theorem	قضیه فوریه
frequency masking	استتار فرکانسی
frequency resolution	وضوح فرکانسی
frequency smearing	محو فرکانسی

functional model	مدل کاربردی
fundamental frequency	فرکانس پایه
G	
Ganong effect	اثر گانونگ
gestalt	گشتالت
gesturalist	بستارگرایی
gesture	بستار
glide	غلت
glottal impedance	مقاومت چاکنایی
glottalized	چاکنایی شده
gradient	مدرج
greatest common denominator	بزرگ‌ترین مقسوم‌علیه مشترک
H	
hamming	همینگ
harmonic	همساز
harmonic motion	حرکت هماهنگ
Helmholtz resonator	بازخوان هلمهولتز
Hertz	هرتز
heterorganic	ناهم‌اندام
hierarchical cluster analysis	تحلیل خوشه‌ای سلسله‌مراتبی
high-pass filter	صافی بالاگذر
I	
identification function	تابع شناسایی
identification task	وظیفه شناسایی
imaginary part	بخش موهومی
implosives	مکیده
impulse	تکانه، پالس
incident wave	موج انتشار
ingressive	درون‌سو

intensity	شدت
International Phonetic Association (IPA)	انجمن بین‌المللی آواشناسی (آی‌پی‌ای)
interpolation	اتصال
inter-speaker variability	تنوع بینافردی
intraoral	دهانی
inverse discrete Fourier transform	تبدیل معکوس گسسته فوریه
inverse filtering	فیلتر معکوس
J	
jittery	لرزان
K	
kinetic energy	انرژی جنبشی
L	
lag	پس‌افت
laminar flow	جریان لایه‌ای
landmark	محدوده
larynx lowering	پایین‌رفتگی حنجره
lateral	کناری
lateral approximant	ناسوده کناری
lateral channel	مجرای کناری
line busy	شلوغی خط
linear predictive coding (LPC)	رمزگذاری پیش‌گویانه خطی (ال‌پی‌سی)
lip protrusion	پیش‌آمدگی لب
lip source	منبع لب
locus frequency	فرکانس کانونی
longitudinal wave	موج طولی
loss-less	بی‌افت
lossy	باافت
low-pass filter	فیلتر پایین‌گذر
M	

magnitude	بزرگی
mapping	نگاشت
masking	استتار
McGurk effect	اثر مک‌گرک
McGurking	مک‌گرک کردن
Mel scale	مقیاس مل
micro-pascal (μPa)	میکروپاسکال
misheard	اشتباه شنیداری
misperception	اشتباه درکی
mispronunciation	اشتباه تلفظی
modal voice	واک عادی
modes of vibration	مُد‌های ارتعاش
modulating	مدوله کردن
monophthong	واکه ساده
movement	جاب‌جایی
moving average smoothing	هموارسازی میانگین متحرک
MP3	ام‌پی‌تری
multidimensional scaling (MDS)	مقیاس‌گذاری چندبعدی (ام‌دی‌اس)
N	
narrow-band	نوارباریک
nasal murmur	زمزمه خیشومی
nasal tract	مجرای خیشومی
nasalization	خیشومی‌شدگی
nasalized vowel	واکه خیشومی‌شده
natural log	لگاریتم طبیعی
natural resonant frequency	فرکانس بازخوانی طبیعی
near merger	ادغام جزئی
node	گره
nomogram	نوموگرام

nonconsonantal	ناهمخوانی
noncontinuant	ناپیوسته
nonlinear	غیرخطی
non-nasalized vowel	واکه غیرخیشومی
non-sibilant	غیرصفیری
nonsinusoidal	غیرسینوسی
nonword	ناکلمه
Nyquist frequency	فرکانس نایکویست
O	
obstacle source	منبع مانع
obstacle turbulence	اغتشاش مانعی
offset	پایانه
online	برخط
onset	آغاز
onset licensing	جواز آغاز
open quotient (oq)	ضریب بازی (اُکیو)
oral	دهانی
oral tract	مجرای دهانی
oscilloscope	اسیلوسکوپ
overall amplitude	دامنه کلی
P	
palatalization	کامی‌شدگی
palatoglossus	زبانی‌کامی
palatogram	کام‌نگاشت
particle velocity	سرعت ذره‌ای
Pass-band	نوارگذر
pass band	نوار گذر
passive nasalization	خیشومی‌شدگی بی‌اختیار
peak amplitude	دامنه قله

peak-to-peak amplitude	دامنهٔ قله به قله
perceived loudness	بلندی درکی
perceptual amplitude	دامنهٔ درکی
perceptual distance	فاصلهٔ درکی
perceptual illusion	توهم درکی
perceptual licensing	جواز درکی
perceptual magnet effect	اثر آهن‌ربای درکی
perceptual map	نقشهٔ درکی
perceptual similarity	شباهت درکی
period	دورهٔ تناوب
periodic	متناوب
periodicity	الگوی تناوبی
peripheral auditory system	سیستم شنوایی بیرونی
perturbation theory	نظریهٔ اختلال
phase	فاز
phonation type	نوع واک‌سازی
phoneme restoration	بازگردانی واج
phonetic coherence	انسجام آوایی
pitch period	دورهٔ تناوب زیروبمی
pitch trace	نمودار زیروبمی
pitch-doubling	دوبرابرشدن زیروبمی
pitch-halving	نصف‌شدن زیروبمی
pixel	پیکسل
plain voiced vowel	واکهٔ واک‌دار ساده
polarity change	تغییر قطبیت
posture	بستار
potential energy	انرژی پتانسیل
power spectrum	طیف قدرت
pre-aspirated	پیش‌دمیده

pre-emphasis	پیش تقویت
preferred vowel	واکه مرجح
pre-glottalized	پیش چاکنایی شده
pre-nasalized	پیش خیشومی شده
pressure fluctuation	نوسان فشار
product array	آرایه حاصل ضرب
prototype	نمونه متعارف
psychoacoustics	سایکواکوستیک
pulmonic	ششی
Pure tone	نواخت خالص
push chain	زنجیره فشار
Q	
quantal region	ناحیه ذره‌ای
quantal theory	نظریه ذره‌ای
quantization	کمی سازی
quantization noise	نوفه کمی سازی
R	
rarefaction	رقت
rectangular	مستطیلی
reflected wave	موج انعکاسی
regions of stability	نواحی پایدار
reject band	نوار دفع
relative loudness	بلندی نسبی
release burst	انفجار رهش
resonant frequency	فرکانس بازخوانی
resonate	تشدیدشدن، بازخوان شدن
response function	تابع پاسخ
resting area	مساحت استراحتی
retroflex	برگشتی

rise time	زمان افزایش
rocking motion	حرکت موجی
root mean square (RMS)	ریشه میانگین مربعات (آرام‌اس)
Ruben's tube	لوله روبن
S	
sampling	نمونه‌برداری
sampling rate	نرخ نمونه‌برداری
saturation	اشباع
sawtooth pattern	الگوی دندان‌اره‌ای
sensory input system	سیستم درون‌داد حسی
Shepard's Law	قانون شپارد
shutting stage	مرحله آمادگی
sibilant	صفیری
side branch	شاخه جانبی
side cavity	حفره جانبی
signal clipping	برش سیگنال
signal to noise ratio	نسبت سیگنال به نوفه
silence	سکوت
simple harmonic system	سیستم هماهنگ ساده
simultaneous masking	استتار هم‌زمان
sine wave	موج سینوسی
sinusoidal	سینوسی
skirt	حاشیه
slip of the ear	لغزش گوش
sone	سون
sonorant	رسا
sound wave	موج صوتی
source-filter theory of speech production	نظریه منبع-صافی تولید گفتار
spectral prominence	برجستگی طیفی

spectral smearing	محو‌ی طیفی
spectral tilt	کجی طیفی
spectrograph	طیف‌نگار
speech perception	درک گفتار
speech synthesizer	سنتزکننده گفتار، گفتار ساز
speed of sound	سرعت صوت
spit spike	تکانه بزاقی
spontaneous nasalization	خیشومی‌شدگی خودبه‌خود
standing wave	موج ایستاده
static	ساکن
steady-state	ایستا
stimulus	محرک
stimulus continuum	پیوستار محرکی
stricture	انسداد
strident	تیز
subglottal pressure	فشار زیر حنجره‌ای
subjective loudness	بلندی ذهنی
sublingual cavity	حفره زیرزبانی
surface area	مساحت سطح
syllable coda	پایانه هجا
syllable onset	آغاز هجا
T	
target	هدف
task	وظیفه
temporal masking	استتار زمانی
tenseness	سختی
timbre	طنین
time resolution	وضوح زمانی
time-varying	متغیر با زمان

tone	نواخت
tone glide	نواخت غلتان
tongue bunching	جمع شدن زبان
tongue retraction	عقب کشیدن زبان
tongue retroflexion	برگشت زبان
tongue-root advancement	جلو کشیدن ریشه زبان
trajectory	منحنی گذر
transient	گذرا
transverse wave	موج عرضی
triangulation	مثلث بندی
tube model	مدل لوله ای
turbulence	اغتشاش
turbulent airflow	جریان هوای اغتشاشی
turbulent flow	جریان اغتشاشی
two-tube model	مدل دولوله ای
U	
ultrasound	فراصوت
unaspirated	نادمیده
uncoupled	جفت نشده
undamped	نامیرا
upward spread of masking	گسترش رو به بالای استتار
V	
valley	دره
velocity	سرعت
velocity wave	موج سرعت
velopharyngeal port	دریچه کامی حلقی
vocal fry	جز جز چاکنایی
vocalic	واکه ای
vocalis	وکالیس

voice bar	خط واک
voice onset time (VOT)	زمان آغاز واک (وی‌اُتی)
voicing	واک
voicing quanta	مقادیر ذره ای واک
vortice	جریان حلقوی
vowel harmony	هماهنگی واکه‌ای
vowel shift	تغییر واکه‌ای
W	
wall source	منبع دیوار
warped	تغییر شکل
wave shape	شکل موج
waveform window	پنجره شکل موج
wavelength	طول موج
whisper	نجوا
wide-band	نوارپهن
window	پنجره
windowing	پنجره‌گذاری
X	
x- ray tracing	ردیابی اشعه X
Z	
zero padding	صفرگذاری

نمایه

انفجاری، ۳۴	آ
انواع واک‌سازی، ۱۰۱، ۲۴۶، ۲۶۳	آهن‌ربای درکی، ۱۸۳، ۳۱۱، ۳۳۴
اوهالا، ۴۵، ۲۵۲، ۲۶۴	
ایسلندی، ۲۸۲، ۲۹۲	۱
ب	
بازخوان، ۱۴، ۵۲، ۵۳، ۵۵، ۵۶، ۵۸، ۶۰، ۶۱	اثر تقابل شنیداری، ۱۵۳
۶۴، ۶۵، ۶۶، ۶۸، ۶۹، ۱۰۷، ۱۹۲، ۱۹۴،	اثر مک‌گرک، ۱۸۲، ۳۱۱، ۳۳۲
۱۹۵، ۱۹۶، ۱۹۷، ۲۰۱، ۲۱۶، ۲۲۹، ۲۳۳،	ارتفاع واکه، ۲۰۷، ۲۹۳
۲۳۴، ۲۵۱، ۲۵۴، ۲۵۶، ۲۶۹، ۲۷۴، ۲۷۵،	استیونس، ۱۰، ۴۳، ۴۷، ۴۹، ۷۳، ۱۲۶، ۱۹۰،
۲۷۶، ۲۸۳، ۲۸۷، ۲۹۴، ۳۰۷، ۳۱۰، ۳۱۲،	۲۰۳، ۲۲۳، ۲۲۸
۳۲۸، ۳۳۰	اشتباه تلفظی، ۳۱۱، ۳۳۲
بازخوان حفره جلوبی، ۱۹۲، ۱۹۷	اشتباه درکی، ۳۱۱، ۳۳۲
بازخوان هلمهولتز، ۱۹۵، ۱۹۶، ۲۱۶، ۳۱۲، ۳۳۰	اصلاح طول، ۶۵، ۶۸، ۷۱، ۷۴، ۳۰۴، ۳۱۱،
بازگردانی واج، ۱۸۴، ۳۱۲، ۳۳۴	۳۲۹
بازنمایی شنیداری، ۶، ۱۳، ۲۰۹، ۲۱۴، ۲۱۵	اغتشاش، ۷، ۲۱۹، ۲۲۰، ۲۲۲، ۲۲۳، ۲۲۴، ۲۲۶،
بدنمایی، ۳۱۳، ۳۲۵	۲۲۹، ۲۴۲، ۲۴۸، ۳۱۲، ۳۲۷، ۳۳۳، ۳۳۸
برش سیگنال، ۳۱۳، ۳۳۶	اغتشاش مانعی، ۲۲۰، ۲۴۲، ۳۱۲، ۳۳۳
بست انسدادی، ۲۴۹، ۲۵۷	اغتشاش مجرای، ۲۲۰، ۲۴۲، ۳۱۲، ۳۲۷
بلندی، ۶، ۴۸، ۴۹، ۶۳، ۱۲۵، ۱۲۶، ۱۲۷، ۱۲۸،	افتاده، ۲۰۷، ۲۱۶، ۲۴۶، ۲۵۴، ۲۶۱، ۲۸۹، ۲۹۲،
۱۲۹، ۱۳۰، ۱۳۱، ۱۳۶، ۱۳۸، ۱۴۰، ۱۴۴،	۳۰۷
۱۴۵، ۱۴۶، ۳۰۵، ۳۱۳، ۳۳۴، ۳۳۵، ۳۳۷	افراشته، ۱۹۷، ۲۰۷، ۲۱۶، ۲۸۸
پ	المان، ۱۶۹
پادبازخوان، ۲۷۴، ۲۷۵، ۲۸۳، ۲۸۶، ۲۹۴، ۳۱۳،	امواج نامتناوب، ۳۳، ۳۴، ۴۱
۳۲۶	انتشار صوت، ۵، ۲۰، ۲۱، ۲۲، ۲۲۹
	انسایشی، ۱۰، ۲۵۸، ۳۰۸، ۳۰۹
	انسجام آوایی، ۳۱۲، ۳۳۴
	انسدادی چاکنایی، ۴۸، ۴۹

- پادسازه، ۲۷۴، ۲۷۵، ۲۷۶، ۲۷۸، ۲۷۹، ۲۸۰، ۲۸۱، ۲۸۳، ۲۸۴، ۲۸۵، ۲۸۶، ۲۸۷، ۲۸۸، ۲۹۲، ۲۹۴، ۳۱۰، ۳۱۳، ۳۲۶
- پادگره، ۵۸، ۷۲، ۷۴، ۲۱۶، ۳۱۳، ۳۲۶
- پالس، ۳۵، ۴۱، ۹۸، ۱۸۰، ۲۴۶، ۳۱۵، ۳۳۰
- پایین رفتگی حنجره، ۳۱۳، ۳۳۱
- پراکنده، ۳۱۳، ۳۲۸
- پرده گوش، ۱۹، ۲۰، ۴۱، ۱۲۴، ۱۴۵
- پسین، ۲۰۷، ۲۱۳، ۲۱۶
- پنجره شکل موج، ۳۱۴، ۳۳۹
- پنجره گذاری، ۳۱۴، ۳۳۹
- پهنای باند، ۳۹، ۴۲، ۱۴۴، ۳۱۴، ۳۲۹
- پیش تقویت، ۳۱۴، ۳۲۹، ۳۳۵
- پیشین، ۱۶، ۴۵، ۱۹۷، ۲۰۷، ۲۱۳، ۲۱۶، ۲۷۲، ۲۸۴
- ت
- تابع تشخیص، ۳۱۴، ۳۲۸
- تابع شناسایی، ۳۱۴، ۳۳۰
- تایلندی، ۲۱۱، ۲۱۲، ۲۷۲، ۲۷۳، ۲۷۸، ۲۷۹، ۲۸۰، ۲۸۳، ۲۸۵
- تبدیل سریع فوریه (افافتی)، ۳۱۴، ۳۲۹
- تجزیه ال پی سی، ۱۵، ۱۱۰، ۱۴۰، ۲۷۶، ۲۷۸
- تجزیه فوریه، ۹، ۹۲، ۱۲۴، ۱۲۵، ۱۳۱، ۳۱۴، ۳۲۹
- تحلیل خوشه ای سلسله مراتبی، ۳۱۴، ۳۳۰
- تحلیل میانگین متحرک خودبازگشتی، ۲۷۶، ۳۱۴، ۳۲۶
- تراکم، ۵۳، ۵۴، ۵۵، ۶۱، ۱۹۸، ۲۵۲، ۲۶۸، ۳۱۴، ۳۲۷
- تغییر واکه ای، ۲۸۸، ۳۱۵، ۳۳۹
- ج
- جبران هم تولیدی، ۱۸۳، ۳۱۵، ۳۲۷
- جریان لایه ای، ۳۱۵، ۳۳۱
- جفت شدگی آکوستیکی، ۱۹۳، ۲۷۵، ۳۰۸
- جواز درکی، ۳۱۵، ۳۳۴
- چ
- چاکنا، ۴۳، ۴۸، ۴۹، ۶۶، ۶۷، ۱۹۰، ۱۹۶، ۱۹۷، ۱۹۸، ۲۰۲، ۲۰۳، ۲۶۹، ۲۷۱، ۲۸۳، ۲۹۴
- ح
- حفره زیربانی، ۳۱۵، ۳۳۷
- حلزون شنوایی، ۱۲۴، ۱۳۶، ۳۱۶، ۳۲۷
- حلزون نگاشت، ۱۴۱، ۱۴۳، ۱۴۵، ۲۱۷، ۲۵۸، ۳۱۶، ۳۲۷
- خ
- خط واک، ۳۱۶، ۳۳۹
- خوسا، ۱۴۰، ۱۴۲
- خیشومی شدگی بی اختیار، ۳۱۶، ۳۳۳
- د
- دامنه قله، ۲۸، ۳۹، ۹۰، ۹۱، ۱۲۲، ۱۲۷، ۳۱۶، ۳۳۳، ۳۳۴
- دامنه قله به قله، ۳۱۶، ۳۳۴
- درک دوطرفه، ۱۶۰، ۱۸۳، ۳۱۶، ۳۲۹
- درک گفتار، ۱۰، ۱۳، ۱۴، ۱۵، ۱۳۵، ۱۳۷، ۱۵۸، ۱۸۰، ۱۸۲، ۱۸۳، ۳۰۶، ۳۱۶، ۳۳۷
- درک مقوله ای، ۱۵۶، ۱۸۳، ۳۱۶، ۳۲۷
- دسی بل، ۳۹، ۹۳، ۹۶، ۱۰۸، ۱۲۷، ۱۲۸، ۱۴۵، ۲۴۷، ۳۰۴

- دمش، ۱۸۳، ۲۴۷، ۲۴۸، ۲۴۹، ۲۵۰، ۲۵۴، ۲۵۷، ۲۶۴، ۳۱۶، ۳۲۶
- دوبرابر شدن زیروبمی، ۱۰۲، ۳۱۶، ۳۳۴
- ر
- رسا، ۱۱، ۳۱۷، ۳۳۶
- رسانای آکوستیکی، ۱۹۱، ۳۱۷، ۳۲۵
- رقت، ۲۱، ۲۲، ۴۱، ۵۳، ۵۵، ۳۱۷، ۳۳۵
- رهش انسدادی، ۱۳۴، ۲۴۳، ۲۵۶، ۲۶۴، ۲۶۵
- ریشه میانگین مربعات (آرام‌اس)، ۳۱۷، ۳۳۶
- ز
- زمان افزایش، ۲۵۸، ۲۶۴، ۲۶۵، ۳۱۷، ۳۳۶
- زمان آغاز واک (وی‌آتی)، ۱۴۹، ۳۱۷، ۳۳۹
- زمزمه خیشومی، ۲۷۳، ۲۷۹، ۲۸۰، ۲۸۱، ۳۱۷، ۳۳۲
- س
- سازه خیشومی، ۲۷۹، ۲۸۷
- سازه، ۶۸، ۷۴، ۱۱۶، ۲۰۳، ۲۱۷، ۲۵۴، ۲۵۶
- ۲۷۲، ۲۷۳، ۲۷۸، ۲۸۳، ۲۸۸، ۲۸۹، ۲۹۲
- ۳۱۷، ۳۲۰، ۳۲۹
- سرعت صوت، ۶۰، ۶۱، ۶۷، ۷۴، ۱۹۱
- ۱۹۴، ۲۷۱، ۳۱۷، ۳۳۷
- سطح مقطع عرضی، ۱۸۹، ۱۹۰، ۱۹۳، ۱۹۴
- ۲۰۱، ۲۱۶
- سون، ۱۲۵، ۱۲۷، ۱۲۹، ۱۴۴، ۱۴۵، ۱۴۶، ۳۰۵
- ۳۱۷، ۳۳۶
- سیستم شنوایی بیرونی، ۶، ۱۲۳، ۱۲۴، ۱۳۳
- ۱۴۰، ۱۴۵، ۳۱۸، ۳۳۴
- سیستم هماهنگ ساده، ۵۸، ۳۱۸، ۳۳۶
- سیکل، ۲۶، ۴۴، ۵۲، ۵۴، ۵۵، ۵۹، ۶۰، ۶۶
- ۱۰۱، ۱۰۲، ۲۴۵، ۳۱۸، ۳۲۸
- سیگنال پیوسته، ۳۱۸، ۳۲۸
- سیگنال دیجیتال، ۹، ۷۸
- سیگنال گسسته، ۳۱۸، ۳۲۸
- سینوس، ۲۷۸، ۲۹۴
- ش
- شبهات درکی، ۱۰، ۳۱۸، ۳۳۴
- شپارد، ۱۷۵، ۱۷۶، ۱۷۷، ۳۱۹، ۳۳۶
- شیب صافی، ۳۱۸، ۳۲۹
- ص
- صافی، ۱۰۲
- صافی بالاگذر، ۱۰۸، ۳۱۸، ۳۳۰
- صافی پادبندنمایی، ۳۱۸، ۳۲۶
- صافی پایین‌گذر، ۱۰۳، ۳۱۹، ۳۳۱
- صافی دستگاه گفتار، ۵، ۷، ۳۹، ۴۳، ۴۷، ۵۰
- ۷۰، ۱۱۱، ۲۱۹، ۲۲۲، ۲۲۶، ۲۵۰، ۲۵۳
- ۲۶۷
- صافی نوارگذر، ۳۱۸، ۳۲۶
- صفرگذاری، ۱۱۴، ۳۱۸، ۳۳۹
- ض
- ضرایب ال‌پی‌سی، ۱۱۱، ۱۱۳، ۱۱۴
- ضریب بازی، ۲۴۵، ۲۴۷، ۳۱۸، ۳۳۳
- ط
- طول موج، ۶۸، ۶۹، ۷۴
- طیف اف‌اف‌تی، ۹۶، ۱۰۵، ۱۱۴، ۲۷۷، ۲۷۸
- ۲۸۳، ۲۸۵، ۳۰۵
- طیف آکوستیکی، ۱۳۹، ۱۴۰، ۲۰۸

- طیف شنیداری، ۱۰، ۱۳۸، ۱۴۰، ۱۴۲، ۱۴۵، ۲۱۰، ۳۰۷، ۳۱۹، ۳۲۶
- طیف قدرت، ۴۴، ۱۱۶، ۱۱۷، ۱۴۲، ۲۲۲، ۲۴۷، ۲۷۰، ۳۱۹، ۳۳۴
- طیف‌نگار، ۲۰۷، ۳۱۹، ۳۳۷
- طیف‌نگاشت، ۹، ۱۱۶، ۱۱۷، ۱۲۰، ۱۴۱، ۱۶۲، ۱۶۸، ۱۹۱، ۲۰۹، ۲۴۸، ۲۷۲، ۲۷۳، ۲۷۸، ۲۸۰، ۲۸۵، ۲۸۸
- ف**
- فاز، ۲۶، ۳۲، ۳۳، ۴۱، ۶۰، ۷۶، ۷۹، ۸۴، ۲۶۸، ۲۷۶، ۳۱۹، ۳۳۴
- فاصله درکی، ۱۸۴، ۳۰۶، ۳۱۹، ۳۳۴
- فانت، ۱۸۹، ۲۱۴، ۲۷۱، ۲۷۵، ۲۸۱، ۲۸۲
- فرکانس بازخوانی، ۵۲، ۵۶، ۵۸، ۵۹، ۶۱، ۶۳، ۶۴، ۶۵، ۶۷، ۶۸، ۷۲، ۷۴، ۱۲۶، ۱۴۵، ۱۹۵، ۱۹۷، ۱۹۸، ۲۷۱، ۲۷۳، ۲۸۶، ۳۱۹، ۳۳۲
- فرکانس پایه، ۳۳۰، ۳۱۹، ۳۴۶
- فرکانس شنیداری، ۱۳۲، ۱۳۳، ۱۳۹، ۱۴۰، ۱۴۵
- فرکانس کانونی، ۳۰۸، ۳۰۹، ۳۱۹، ۳۳۱
- فشار بالای هوای دهانی، ۲۵۲
- فشار زیرحنجره‌ای، ۳۱۹، ۳۳۷
- فشرده، ۲۰، ۶۳، ۱۹۸، ۲۲۹، ۲۴۲، ۳۱۷، ۳۱۹، ۳۲۷
- فضای واکه‌ای، ۶، ۲۰۵، ۲۸۸، ۳۱۹، ۳۲۵
- فوجیمورا، ۲۷۵، ۲۸۴
- فیلتر معکوس، ۲۷۶، ۲۷۷، ۳۱۹، ۳۳۱
- گ**
- کامی‌شدگی، ۲۶۱، ۳۲۰، ۳۳۳
- کانتونی، ۱۱۷، ۱۴۱، ۱۴۳
- کانون سازه‌ای، ۲۶۳
- کدگذاری پیشگویانه خطی (ال پی سی)، ۹۸، ۱۰۷، ۳۱۷، ۳۳۱
- کمی‌سازی، ۳۲۰، ۳۲۳
- کویر، ۱۸۲، ۳۱۹، ۳۲۸
- ک**
- گانوگ، ۱۶۶، ۱۶۷، ۱۸۲، ۱۸۵، ۳۱۱، ۳۳۰
- گذرا، ۳۴، ۳۵، ۴۱، ۸۹، ۹۳، ۱۸۰، ۳۲۰، ۳۳۸
- گذرهای سازه‌ای، ۲۵۳، ۲۸۱، ۲۹۰، ۲۹۳، ۳۰۸
- گردی لب، ۷۴، ۲۰۲، ۲۰۴
- گره، ۴۶، ۵۸، ۷۲، ۷۴، ۲۱۶، ۳۲۰، ۳۳۲
- گوش میانی، ۱۲۴، ۱۳۶، ۱۴۵
- ل**
- لاینز، ۱۰۲، ۱۴۱، ۱۴۳، ۱۴۴
- لغزش‌های گوش، ۱۶۶، ۳۲۰، ۳۳۶
- لیندبلوم، ۱۳۸، ۲۰۹
- لیندو، ۲۰۲، ۲۰۸
- م**
- ماتریس ابهام، ۱۷۸، ۱۸۴، ۳۲۰، ۳۲۸
- ماراتی، ۲۹۰، ۲۹۱
- مالایالام، ۲۹۰، ۲۹۱، ۲۹۲
- مبدل آنالوگ به دیجیتال، ۸۸
- محاسبه زیرویمی به روش خودهمبستگی، ۳۲۰، ۳۲۶
- محل تولید، ۲۰۴، ۲۵۰، ۲۶۳، ۲۷۶، ۲۷۹، ۲۸۱، ۲۹۰، ۲۹۳، ۳۰۸
- محو طیفی، ۱۳۹، ۳۲۱، ۳۳۷
- مدل لوله ای، ۱۰، ۲۰۲، ۲۰۳، ۲۲۴، ۲۲۸، ۳۲۱، ۳۳۸
- مدیسون، ۲۰۴

- مرحله گرایش، ۲۴۴، ۲۴۸، ۲۵۰، ۲۵۳، ۲۵۴، ۲۵۶
- مزاتک، ۲۰۶، ۲۰۷، ۲۰۹، ۲۱۷، ۳۰۷
- مقاومت آیرودینامیکی، ۳۲۱، ۳۲۵
- مقاومت چاکنایی، ۳۲۱، ۳۳۰
- مقیاس بارک، ۱۳۲، ۱۳۳، ۱۴۴، ۱۴۵، ۲۱۵، ۳۲۱، ۳۲۶
- مقیاس گذاری چندبعدی، ۱۷۹، ۳۲۱، ۳۳۲
- منحنی بلندی همسان، ۱۲۸، ۱۳۰، ۱۳۸، ۱۴۵، ۳۲۲، ۳۲۹
- موج ایستاده، ۵۸، ۵۹، ۶۱، ۶۲، ۶۳، ۶۶، ۶۷، ۶۸، ۷۱، ۷۴، ۱۹۸، ۲۰۰، ۳۲۲، ۳۳۷
- موج سینوسی، ۵۳، ۵۴، ۵۵، ۵۷، ۵۸، ۵۹، ۶۰، ۶۱، ۶۳، ۷۱، ۸۴، ۹۳، ۱۱۶، ۱۳۶، ۱۴۵، ۲۶۷، ۲۶۸، ۲۶۹، ۳۲۲، ۳۳۶
- موج متناوب ساده، ۲۵، ۴۱
- موج متناوب مرکب، ۴۴، ۴۷، ۵۰
- میرایی، میرا، ۲۶۸، ۲۶۹، ۲۷۰، ۲۹۴، ۳۰۹
- ن**
- ناپیوسته، ۳۲۲، ۳۳۳
- ناهمخوانی، ۳۲۲، ۳۳۳
- نچ آوا، ۹۲
- نسبت سیگنال به نوفه، ۳۲۳، ۳۳۶
- نصف شدن زیرویمی، ۱۰۱
- نظریه اختلال، ۱۹۷، ۱۹۹، ۲۰۱، ۲۰۲، ۲۰۳، ۲۱۶، ۲۷۱، ۳۲۳، ۳۳۴
- نظریه پراکندگی تطبیقی، ۳۲۳، ۳۲۵
- نظریه ذره‌ای، ۱۰، ۴۷، ۷۳، ۳۲۳، ۳۳۵
- نظریه منبع-صافی، ۱۰، ۴۳، ۳۲۳، ۳۳۶
- نقشه درکی، ۱۰، ۱۷۹، ۲۸۹، ۳۲۳، ۳۳۴
- نمونه برداری، ۱۱۶، ۳۰۵، ۳۲۲، ۳۲۳، ۳۳۶
- نواحی پایدار، ۳۰۸، ۳۲۳، ۳۳۵
- نواحی ذره‌ای، ۲۱۵، ۳۰۸
- نوار دفع، ۳۸، ۱۰۵، ۳۰۳، ۳۳۵
- نوار گذر، ۳۸، ۳۹، ۱۰۵
- نوفه سفید، ۳۳، ۳۴، ۳۵، ۴۱، ۸۷، ۲۲۱
- نوفه کمی سازی، ۸۶، ۸۷، ۸۸، ۸۹، ۱۲۱
- نوموگرام، ۱۹۲، ۱۹۶، ۲۰۲، ۲۰۳، ۲۱۶، ۲۳۲، ۲۳۳، ۲۴۲، ۳۲۳، ۳۳۲
- و**
- واک، ۵، ۴۳، ۴۴، ۴۵، ۴۶، ۴۷، ۴۹، ۵۰، ۵۱، ۵۲، ۷۴، ۱۰۱، ۱۰۲، ۱۱۱، ۱۱۵، ۲۴۵، ۲۴۷، ۲۴۸، ۲۶۳، ۳۰۷، ۳۱۰، ۳۱۷، ۳۲۱، ۳۲۳، ۳۲۴، ۳۲۷، ۳۲۸، ۳۳۲، ۳۳۹
- واک بازداشته، ۴۵، ۱۰۱، ۲۴۸، ۳۲۴، ۳۲۸
- واک عادی، ۲۴۵، ۲۴۸، ۳۱۰، ۳۲۴، ۳۳۲
- واک نفسی، ۴۵، ۳۲۴، ۳۲۷
- واکه خیشومی شده، ۳۱۰، ۳۲۴، ۳۳۲
- واکه مرکب، ۱۱۸
- ه**
- همساز، ۴۵، ۴۷، ۵۱، ۷۴، ۱۰۷، ۱۱۰، ۱۱۴، ۱۱۵، ۲۰۹، ۲۱۰، ۲۴۶، ۳۰۷، ۳۲۴، ۳۳۰

ACOUSTIC and AUDITORY PHONETICS

KEITH JOHNSON

Translated by

ZAHRA MAHMOODZADEH

Iranian Research Institute for Information Science and Technology (IRANDOC)

and

RAHIMEH ROOHPARVAR

Shahid Bahonar University of Kerman



**Iranian Research Institute for
Information Science and Technology
(IRANDOC)**



**Chapar
Publication**



آواشناسی آکوستیک و شنیداری

نویسنده: کیت جانسون
ترجمه: دکتر زهرا محمودزاده
دکتر رحیم روحپاروار

شابک: ۹۷۸-۹۶۴-۷۵۱۹-۹۳-۹

ISBN: 978-964-7519-93-9

ACOUSTIC and AUDITORY PHONETICS

KEITH JOHNSON

Translated by
ZAHRA MAHMOODZADEH
and
RAHIME ROUHPARVAR

ISBN: 978-964-7519-93-9



9 789647 519939



Chapar



Iranian Research Institute for
Information Science and Technology