



Digital Speech Processing

به نام خدا



گروه هوش مصنوعی، دانشکده مهندسی  
کامپیوتر

# گفتار پردازش رقمی

## مقدمه

---

حمیدرضا برادران کاشانی



# سرفصل مطالب

- ❖ معرفی گفتار
- ❖ شاخه های پردازش گفتار
- ❖ کاربردهای پردازش گفتار
- ❖ علوم مرتبط با پردازش گفتار
- ❖ تاریخچه پردازش خودکار گفتار
- ❖ زنجیره گفتار

*Hamidreza Baradaran Kashani*

## چرا گفتار؟

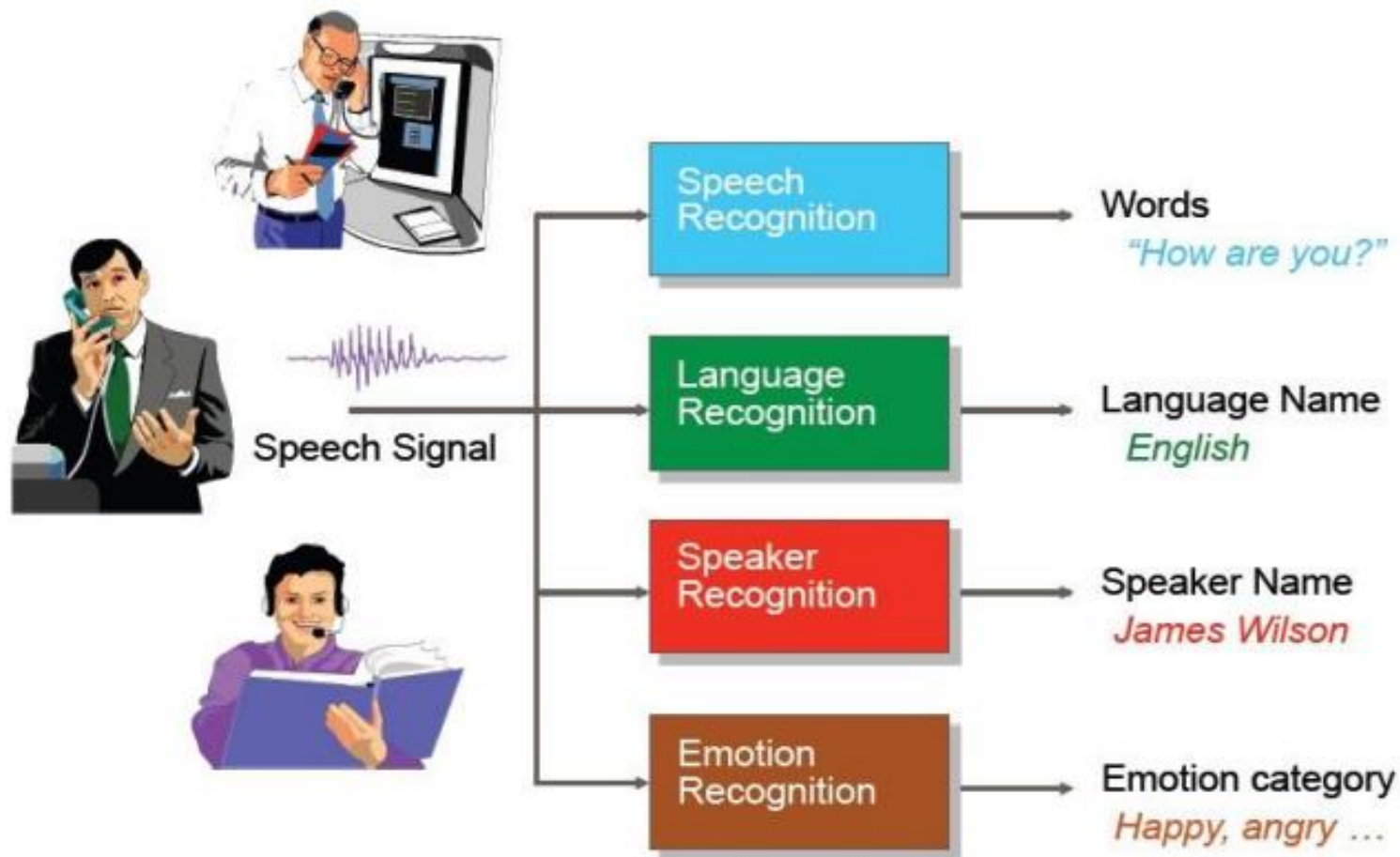
- ۱- ساده ترین راه ارتباطی انسان ها با یکدیگر که بسیار کم هزینه است!
- ۲ - علاوه بر کلمات، مفاهیم و اطلاعات سطح بالاتری را هم منتقل می کند.
- ۳ - نرخ انتقال داده بالایی دارد.
- ۴ - طبیعی است یعنی نیاز به یادگیری ندارد.
- ۵ - کامل تر از زبان نوشتاری است.
- ۶ - بواسطه وجود تاکید و لحن در بیان گفتاری می تواند بهتر از نوشتار مقصود گوینده را مشخص می کند

## چرا پردازش گفتار؟

- ۱- ایجاد ماشین هایی با توانایی ها و قابلیت های انسان ها در صحبت کردن، شنیدن و درک کردن گفتار و ...
  - ۲- ایجاد ارتباط میان انسان و ماشین از طریق گفتار
  - ۳- ساخت سیستم های مکالمه گفتار (Spoken Dialog Systems)
- .....



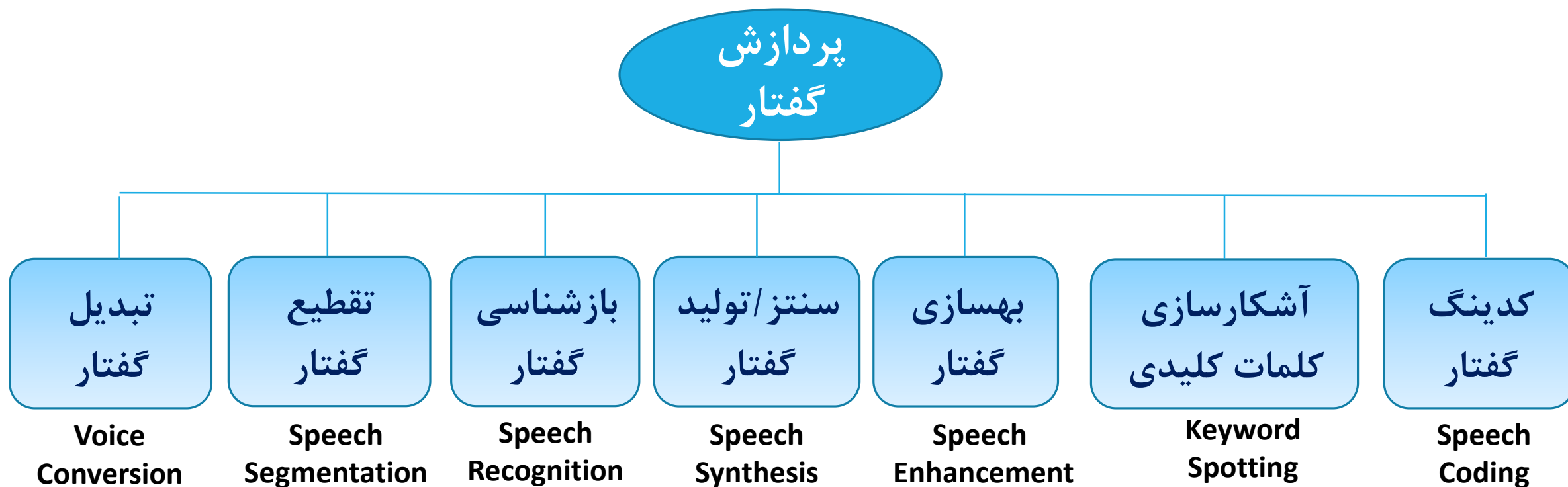
# اطلاعات موجود در گفتار



Hamidreza Baradaran Kashani



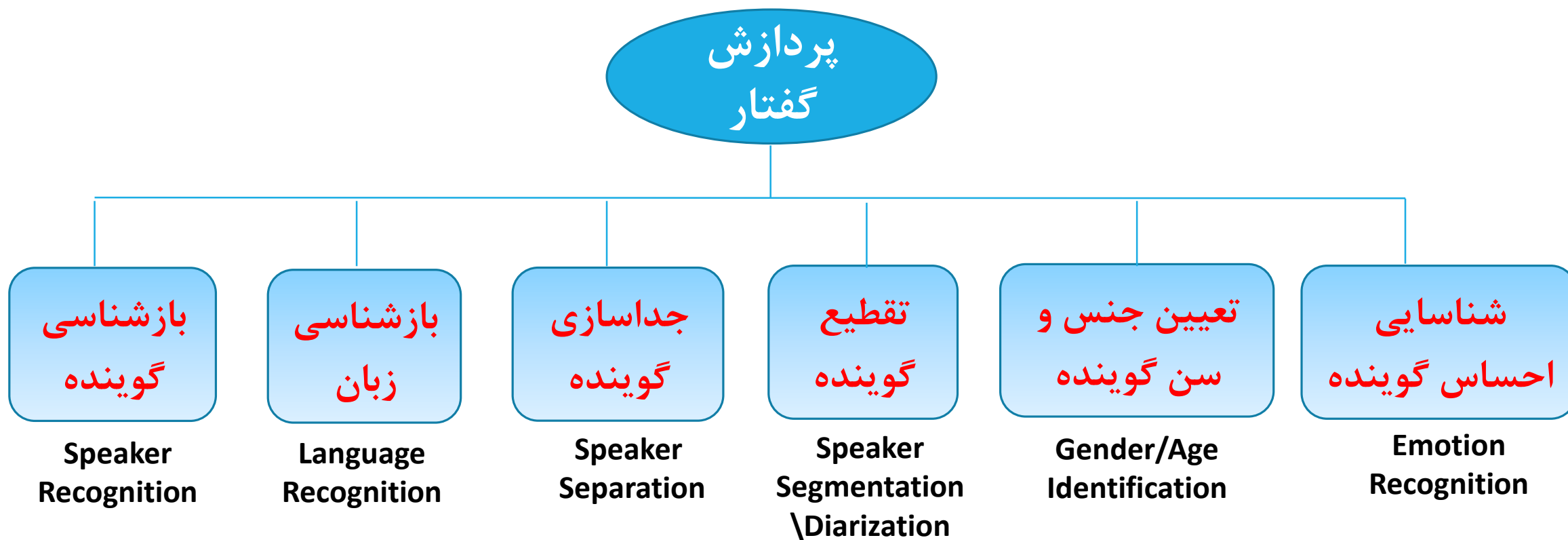
# شاخه های پردازش گفتار



Hamidreza Baradaran Kashani



# شاخه های پردازش گفتار



Hamidreza Baradaran Kashani



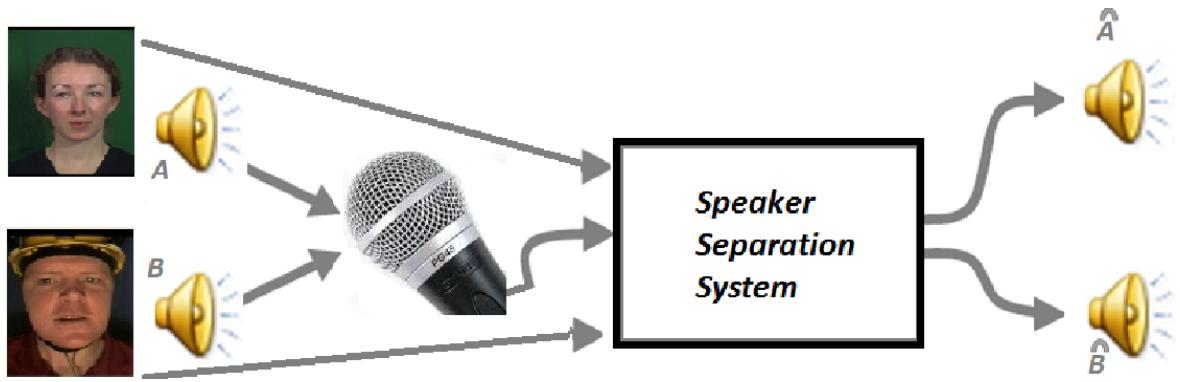
# شاخه های پردازش گفتار



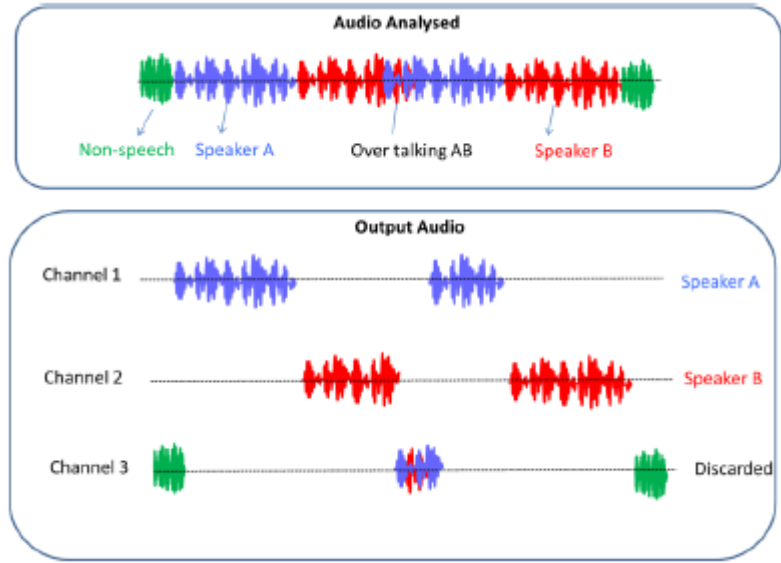
Speech-to-Speech Translation



Voice Conversion



## Speaker Diarization



Hamidreza Baradaran Kashani





# کاربردها



## کاربردهای عمومی

- فرمان های صوتی در اتومبیل، وسایل خانگی و ...
- پردازش گفتار برای کمک به نابینایان،
- آموزش و یادگیری زبان های خارجی (تلفظ صحیح صداها و آهنگ جملات)،
- سیستم های یادگیری بصورت پرسش و پاسخ
- سیستم های اطلاع رسانی مانند تلفن گویا
- کلید و قفل صوتی
- خدمات عمومی مثلا فروش خودکار بلیط در ایستگاه های حمل و نقل
- شناسایی زبان فرد گوینده در سیستم های ترجمه خودکار

Hamidreza Baradaran Kashani

## کاربردهای مخابراتی

- شماره گیری تلفن توسط صدای فرد
- ارسال صدا همراه با تصویر و سایر داده های اطلاعاتی دیگر
- دسترسی از راه دور با پایگاه های داده توسط صوت افراد
- ....



## کاربردهای پزشکی و صنعتی

- تشخیص ناهنجاری ها و بیماری های گفتاری
- ساخت سمعک ها و ابزار کمک شنوایی
- بازشناسی فرامین صوتی در مکان های خاص مثلا کابین هواپیما
- ....

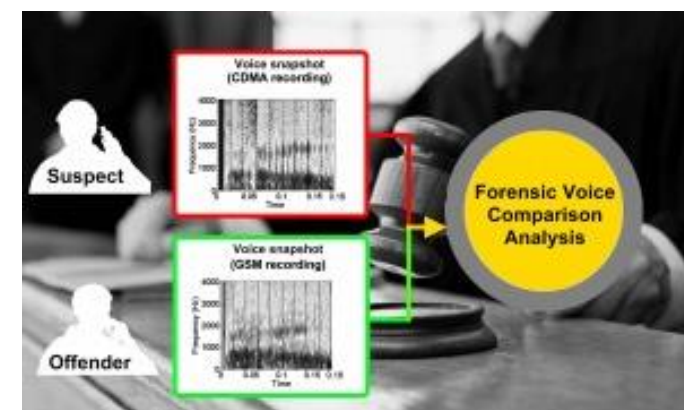
Hamidreza Baradaran Kashani

## کاربردهای اداری

- بازشناسی هویت افراد قبل در هنگام ورود
- دیکته اتوماتیک و بازشناسی گفتار و گوینده در جلسات اداری
- فرامین کنترلی توسط صدا

## کاربردهای امنیتی

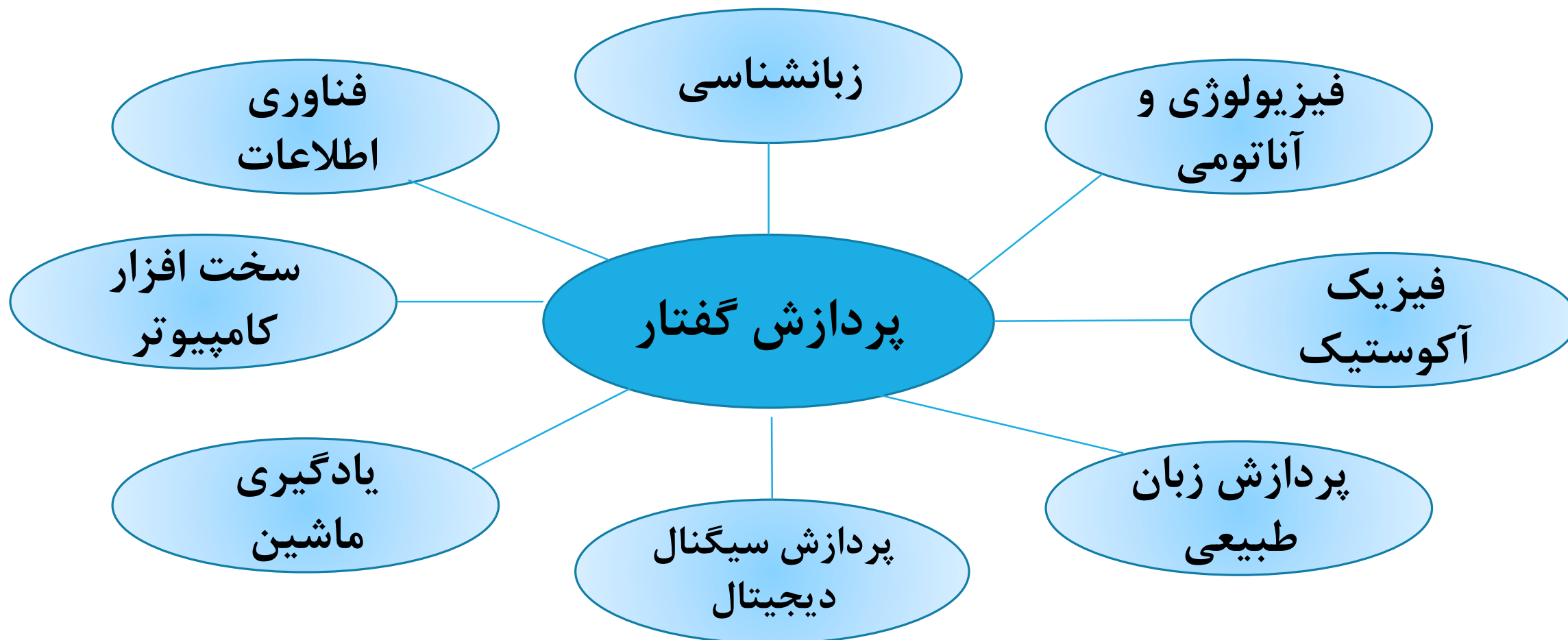
- تشخیص شنود و جاسوسی مبتنی بر گفتار
- فرامین صوتی به ادوات نظامی
- ترجمه زبان مبتنی بر گفتار در مناطق نظامی
- مقایسه صوت فرد مجرم با صوت فرد مظنون به عنوان مدرک در دادگاه



Hamidreza Baradaran Kashani



# علوم مرتبط با پردازش گفتار



Hamidreza Baradaran Kashani



❖ ۱۹۲۰: ساخت اسیلوگراف برای نمایش شکل موج زمانی گفتار

❖ ۱۹۴۶: ساخت اسپکتروگراف صدا در آزمایشگاه Bell

❖ ۱۹۵۸: ساخت یکی از اولین سیستم های بازشناسی گفتار توسط Denis Fry در UCL

❖ ۱۹۷۰ به بعد: ارائه تکنیک های پردازش سیگنال دیجیتال

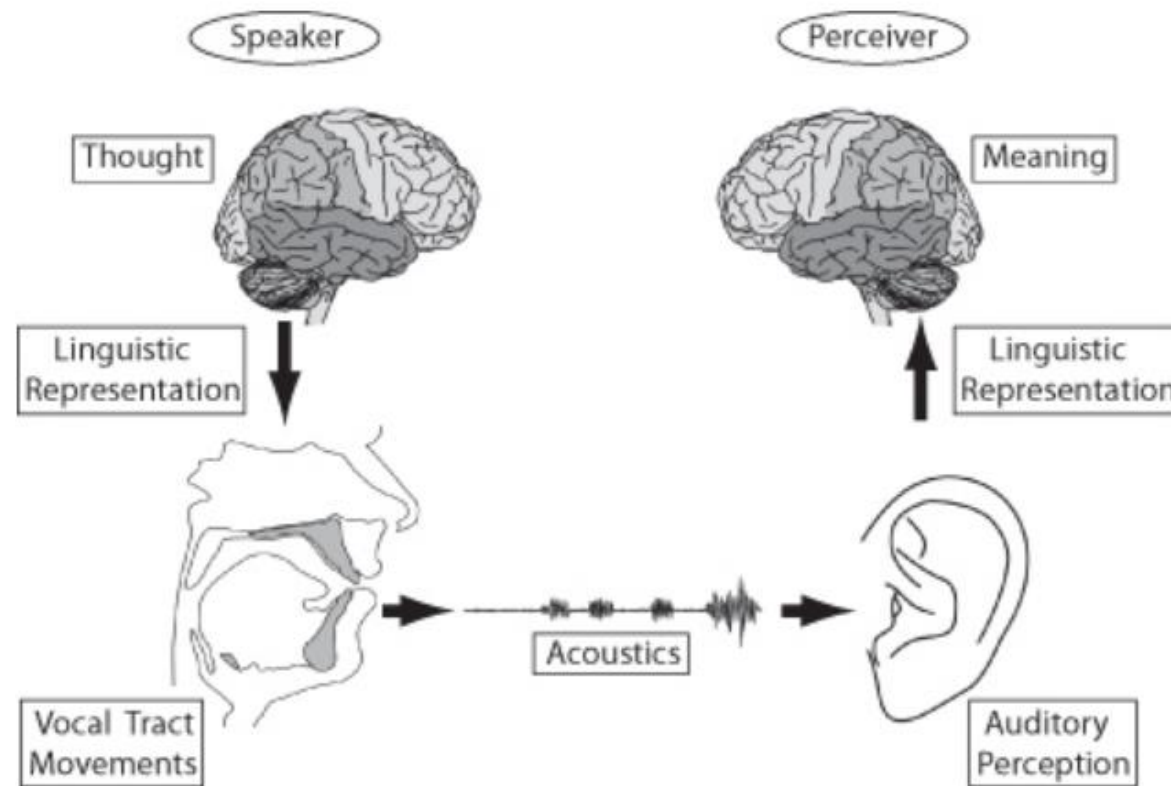
❖ آنالیز پیشگویی خطی (LPC)، تبدیل فوریه سریع (FFT)، استخراج پارامترهای گفتاری مثل گام و فرمنت، آنالیز کپسترال، بکارگیری مدل مخفی مارکوف و رویکردهای مختلف یادگیری ماشین و ....

❖ ۱۹۹۰: ساخت یکی از سیستم های تجاری بازشناسی گفتار توسط آقای Rabiner

❖ از ۲۰۱۳ تا کنون: رویکردهای مختلف یادگیری عمیق در اکثر کاربردهای پردازش گفتار



## *Feed-forward, auditory-only speech chain* (Denes and Pinson, 1993)



(image by W. Murphey and A. Yeung).

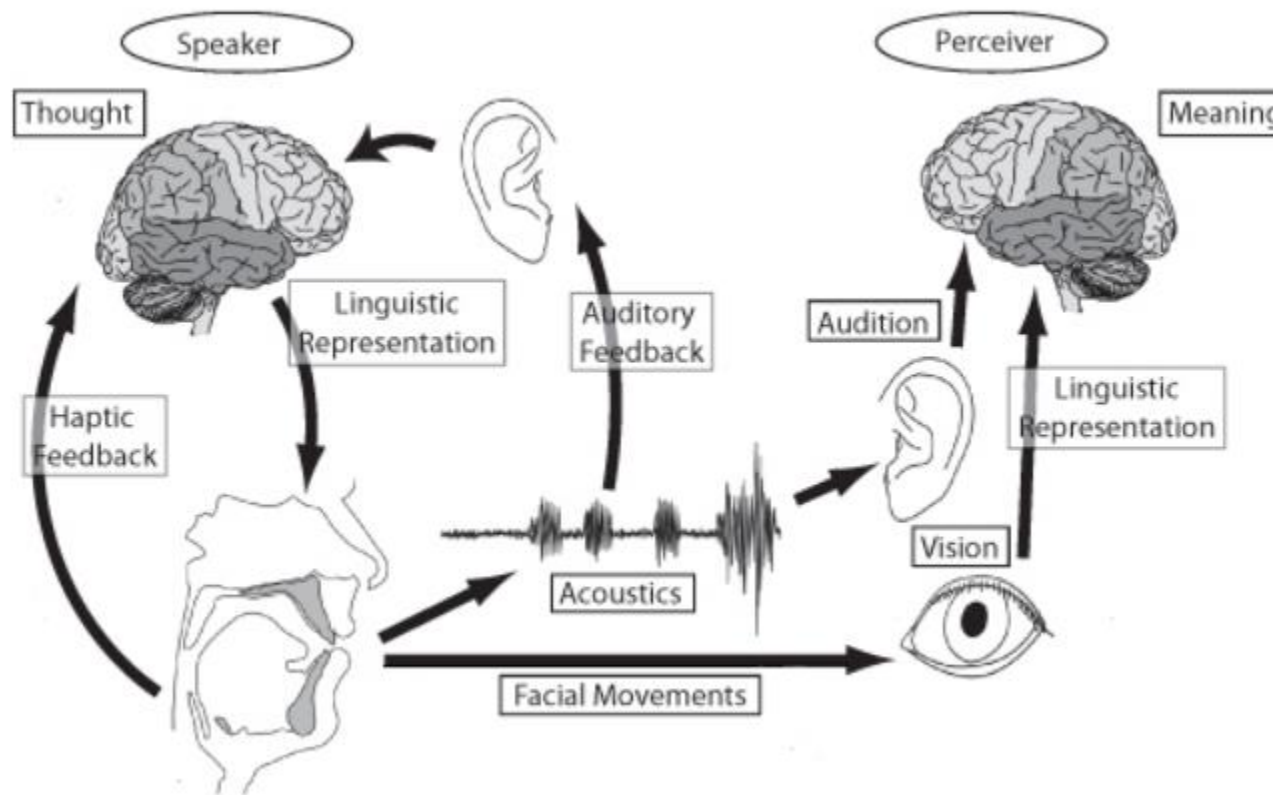
Hamidreza Baradaran Kashani





# زنجیره گفتار

*Multimodal speech chain with feedback loops*  
(Denes and Pinson, 1993)



(image by W. Murphey and A. Yeung).

Hamidreza Baradaran Kashani