

**1- مقدمه**

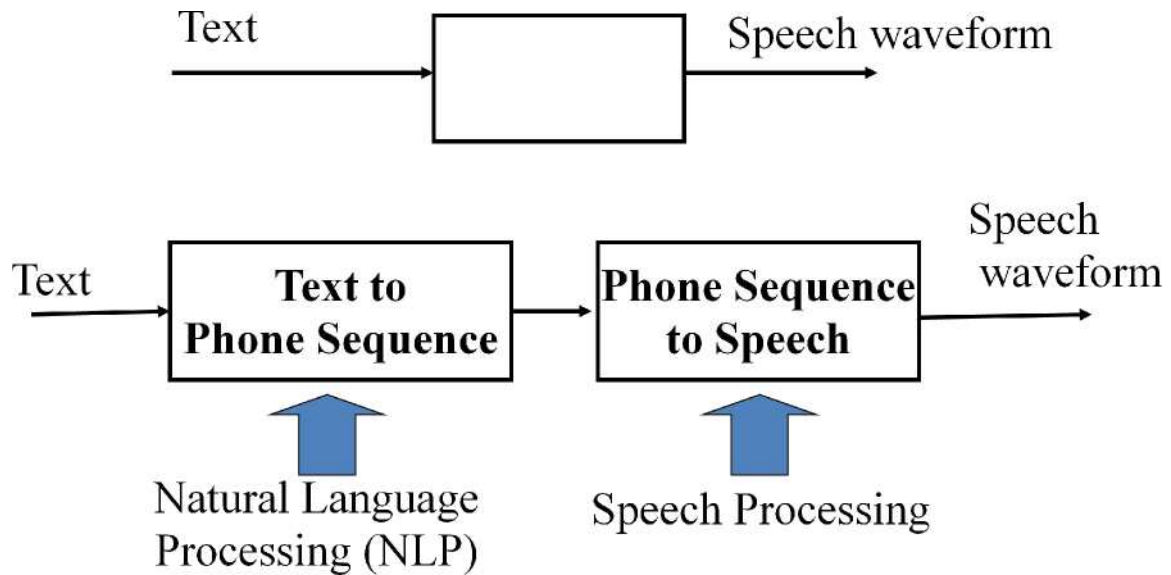
تبدیل متن به گفتار

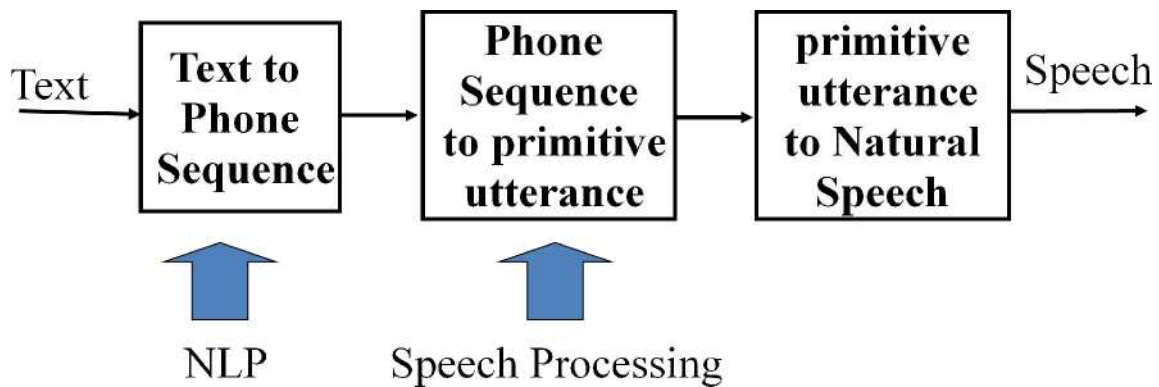
- مبتنی بر قانون
- مفصلی
- چسبانندی (concatenative)
- انتخاب واحد
- آماری (مبتنی بر مدل مخفی مارکوف)

**2- مفاهیم اولیه**

در این فصل بحث تبدیل دنباله آوایی به شکل موج را بررسی می کنیم (تصویر 1).

یعنی فقط بحث پردازش سیگنال تبدیل متن به گفتار را و نه بحث پردازش زبان طبیعی (NLP) بررسی می کنیم.





تصویر 1 - تبدیل دنباله آوایی به شکل موج

برای طبیعی بودن گفتار باید موارد زیر را در نظر بگیریم:

- انرژی گفتار
- مدت تلفظ واج ها
- گام
- آهنگ
- تاکید

آهنگ و تاکید در طبیعی بودن گفتار بسیار موثر هستند.

تعریف آهنگ: تغییر فرکانس گام در حین صحبت کردن

تعریف تاکید: افزایش فرکانس گام در یک زمان مشخص

آهنگ گفتار بستگی به زمینه دارد.

- معمولاً اطلاعات بیان شده در جواب یک سؤال آهنگ بیشتری دارد.
- در حالی که اطلاعاتی که از قبل می دانیم آهنگ دار نیست.

مثال:

سؤال 1: What types of foods are a good source of vitamins?

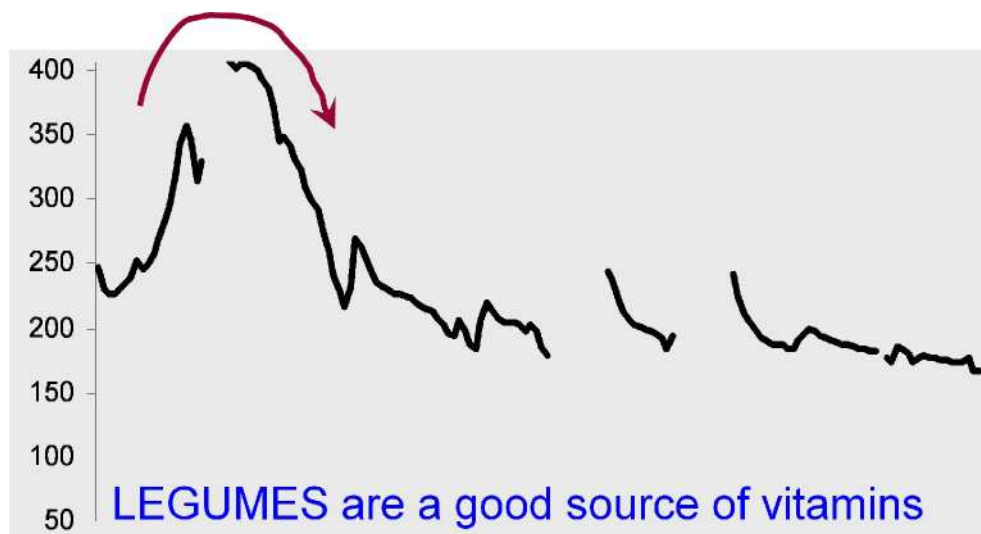
جواب 1: LEGUMES are a good source of vitamins. (تصویر 1)

سؤال 2: Are legumes a source of vitamins?

جواب 2: Legumes are a GOOD source of vitamins. (تصویر 2)

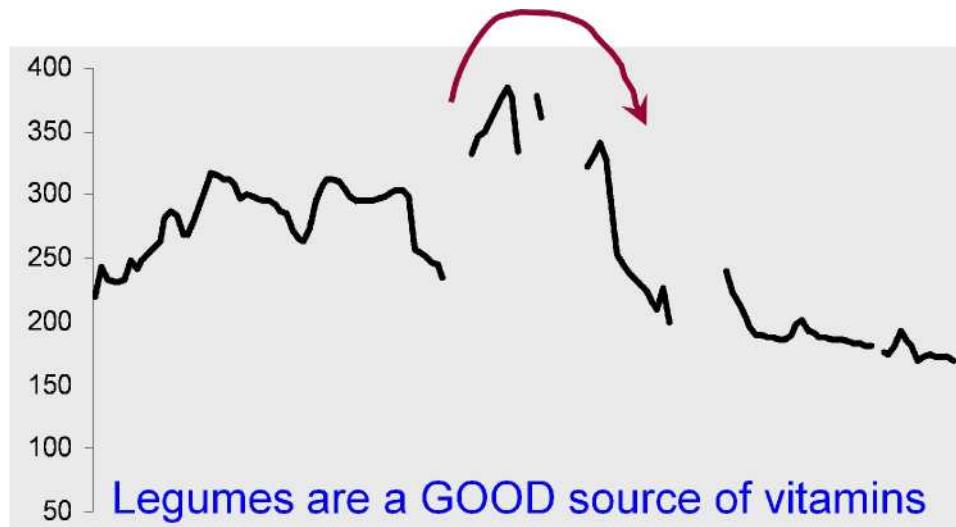
سؤال 3: What are legumes a good source of ?

جواب 3: Legumes are a good source of VITAMINS. (تصویر 3)



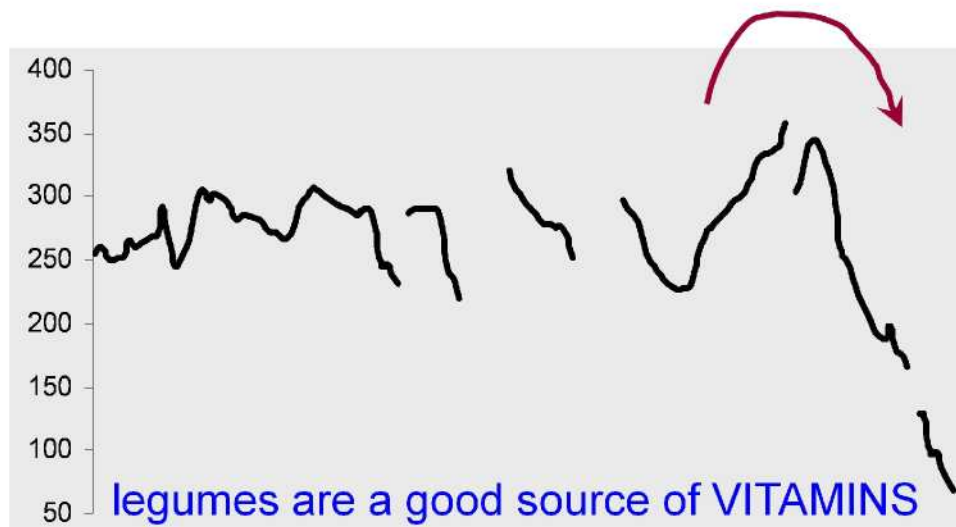
The main **rise-fall** accent (= “I assert this”) shifts locations.

تصویر 1 – جواب 1



The main **rise-fall** accent (= "I assert this") shifts locations.

نص



The main **rise-fall** accent (= "I assert this") shifts locations.

تصویر 3 - جواب 3



همان طور که گفتیم، در این فصل روش های تبدیل دنباله واجی (نه متن) به شکل موج را مشاهده می کنید.

روش های بررسی شده در این فصل عبارتند از:

- سنتز مفصلی (articulatory)
  - حرکت مفاصل، اعضا و ویژگی های صوتی مسیر صوتی انسان را مدل می کند.
- سنتز الحاقی (Concatenative)
  - از دیتابیسی از نمونه گفتارهای ذخیره شده برای تولید شکل موج نهایی استفاده می کند.
  - سنتز دایفون (diphone)
  - سنتز انتخاب واحد (unit selection)
- سنتز آماری (مبتنی بر مدل مخفی مارکوف)
  - پارامترهایی را از روی دیتابیس گفتار آموزش می دهد.
- مبتنی بر قانون (rule-based)
  - استفاده از قوانین و فیلترهایی برای تولید شکل موج

## 2- سنتز مفصلی

شبیه سازی فرآیندهای فیزیکی تولید گفتار انسان

در برخی روش ها از به هم وصل کردن لوله هایی برای ساختن ماشین های مکانیکی سخنگو استفاده شده است.

در روش های جدید تاثیر مکان اعضا، شکل مسیر صوتی و ... بر روی جریان هوا «شبیه سازی» می شود و خروجی نهایی تولید می شود.

## 3- سنتز الحاقی

دو روش اصلی وجود دارد:

1. الحاق واحدهای واجی (diphone concatenation)

- مثال: اتصای نمونه های دایفون ها یا سیلاب های ضبط شده



## 2. انتخاب واحد آوایی (unit selection)

- استفاده از تعداد زیادی نمونه برای هر واحد آوایی و انتخاب بهترین نمونه در هنگام سنتز

در هر دو روش باید واحد آوایی تعریف شود.

در این بخش واحد های آوایی را بررسی می کنیم:

- پاراگراف
- جمله
- کلمه (بستگی به زبان دارد و تعدا کلمات تا 100 هزار کلمه در زبان نیز می باشد)
- سیلاب
- دایفون و ترایفون
- واج (بین 10 تا 100 وابسته به زبان ها)