DAY 27 · 3부 뭘 할 수 있나

폰은 내 말을 어떻게 받아 적을까? — 목소리

하루 5분 AI 수업 · 2026년 10월 3일

목소리는 잘게 쪼갠 떨림이고, 글자는 그 조각들을 앞뒤와 맞춰 고른 답이야.

한 줄로 말하면

1. AI는 네 말을 통째로 알아듣는 게 아니라 **아주 짧은 조각으로 쪼개서** 들어. 소리를 1초에 1만 6천 번 숫자로 받아 적고, 그걸 25밀리초짜리 창으로 훑어 가.

2. 조각 하나하나는 그냥 '높낮이 무늬'야. 어느 글자인지는 **앞뒤 조각을 같이 봐야** 정해져 — 어제 배운 영상처럼, 한 칸만 봐서는 모르거든.

3. 말하는 건 이 길을 거꾸로 가는 거야. 글자에서 소리 무늬를 만들고, 그 무늬대로 **공기의 떨림을 1초에 1만 6천 번 넘게 찍어** 스피커로 내보내.

그림 한 장으로

말 → 글자 → 말, 한 바퀴



가운데 두 칸이 핵심이야. 잘게 쪼개고, 앞뒤를 같이 보는 것.

핵심: 목소리는 잘게 쪼갠 떨림이고, 글자는 그 조각을 앞뒤와 맞춰 고른 결과다

'안녕하세요'가 글자가 됐다가 다시 소리가 되기까지

1. 소리를 조각내면

1초를 아주 짧은 칸으로 자른다

0~25ms | 25~50ms | 50~75ms | 75~100ms | ... | 1초째

조각 하나는 눈 깜빡할 시간의 몇십 분의 일이야. 이 정도로 잘게 잘라야 'ㅅ'과 'ㅈ'처럼 비슷한 소리가 구분돼.

2. 조각을 글자로

앞뒤를 같이 봐야 글자가 정해진다

안 | **녕** | **하** | **세** | **요** | 않

소리만 들으면 '안'과 '않'은 거의 같아. 그래서 AI는 옆 글자와 어울리는지까지 보고 고르지. 그래서 이름이나 처음 듣는...

3. 글자를 다시 소리로

무늬를 만들고 떨림으로 찍는다

글자 | **높낮이 무늬** | **떨림 값** | **스피커** | 듣기 끝

말소리는 결국 공기가 떨린 기록이야. 그 값을 순서대로 찍어 내면 목소리가 되고, 숨소리나 말 끝 억양까지 흉내 낼 수 있어.

버튼을 눌러 단계를 바꿔 봐. 말소리가 조각으로 쪼개지고, 조각이 글자로 붙고, 글자가 다시 떨림으로 돌아가는 과정이야.

웹 수업에서 직접 해 볼 수 있어요 · next-token.net/learn/day-27/

확인해보기

세 문제만 풀어 보자

Q1 AI가 말소리를 글자로 바꿀 때 가장 먼저 하는 일은?

① 아주 짧은 조각으로 쪼갠다
② 문장 전체를 한 번에 알아듣는다
③ 사전에서 낱말을 찾는다
④ 말한 사람 얼굴을 본다

Q2 '안'과 '않'처럼 소리가 비슷할 때 AI는 어떻게 고를까?

① 더 큰 소리 쪽으로 고른다
② 앞뒤 글자와 어울리는지까지 보고 고른다
③ 무조건 짧은 글자로 고른다
④ 사람에게 매번 물어본다

Q3 글자를 목소리로 만들 때 마지막에 나오는 건?

① 사진 한 장
② 낱말 목록
③ 공기의 떨림 값
④ 인터넷 검색어

왜 그런지까지 알고 가자

Q1

① 소리를 1초에 1만 6천 번 숫자로 받아 적고, 25밀리초짜리 조각으로 훑어. 쪼개야 비슷한 소리들이 구분되거든.

Q2

② 조각 하나는 높낮이 무늬일 뿐이야. 옆 글자와 어울리는지 같이 봐야 정해져. 그래서 처음 듣는 이름은 자주 틀려.

Q3

③ 글자에서 높낮이 무늬를 만들고, 그 무늬대로 떨림 값을 1초에 수천 번 찍어 스피커로 내보내. 받아 적기를 거꾸로 간 셈이야.

목소리는 잘게 쪼갠 떨림이고,
글자는 그 조각들을 앞뒤와 맞춰 고른 답이야.



내일은 검색과 뭐가 다를까 — 목록으로 주는 검색과 한 덩어리로 답하는 AI, 뭐가 다른지.