

A ChatGPT hangmodulja (ChatGPT Voice)

Szerző: Lengyel Zsolt, iTStudy Hungary Számítástechnikai Oktató- és Kutatóközpont

Lektor: Hartyányi Mária, iTStudy Hungary Számítástechnikai Oktató- és Kutatóközpont

Tanulási cél

A beszéd a kommunikáció legtermészetesebb formája – és az MI-asszisztensek ma már valós idejű szóbeli beszélgetésre is képesek. A tananyag célja bemutatni a hangalapú MI-interakciót és lehetséges szerepét a tanításban és a tanulásban. A tananyag végére a tanárok meg tudják különböztetni a diktálást a hangalapú beszélgetéstől, le tudják írni, hogyan működik a ChatGPT Voice jelenlegi változata, azonosítani tudják értelmes felhasználási területeit az ötleteléstől, a nyelvgyakorlason, az akadálymentességen és a gyors szakmai segítségen át a reflektív párbeszédig, és figyelembe tudják venni a gyakorlati korlátokat – például az átírási hibákat, a tantermi zajt és az adatvédelmet –, amikor tanulókkal használják a hangmodult.

Bevezetés: a hang kétféle használata

A ChatGPT hanggal két gyakorlati formában használható. A **diktálás** egyszerűen szöveggé alakítja a beszédünket: beszélünk, a rendszer leírja, és úgy küldjük el az üzenetet, mintha begépettük volna. Ez akkor hasznos, ha gyorsan szeretnénk hosszabb gondolatokat megfogalmazni, vagy kényelmetlen a gépelés, például telefonon.

A **hangalapú beszélgetés** valódi párbeszéd: beszélünk az MI-hez, ő pedig szóban válaszol. Ez gyorsabb, természetesebb interakciót tesz lehetővé – főleg ötletelésnél, gyakorlásnál, vagy ha nem akarunk folyamatosan gépelni. A lényeges különbség: a diktálás a beszédből szöveget csinál, a hangalapú beszélgetés viszont párbeszédese felület, ahol az MI válasza is hangban érkezik.

Hogyan működik ma a ChatGPT Voice?

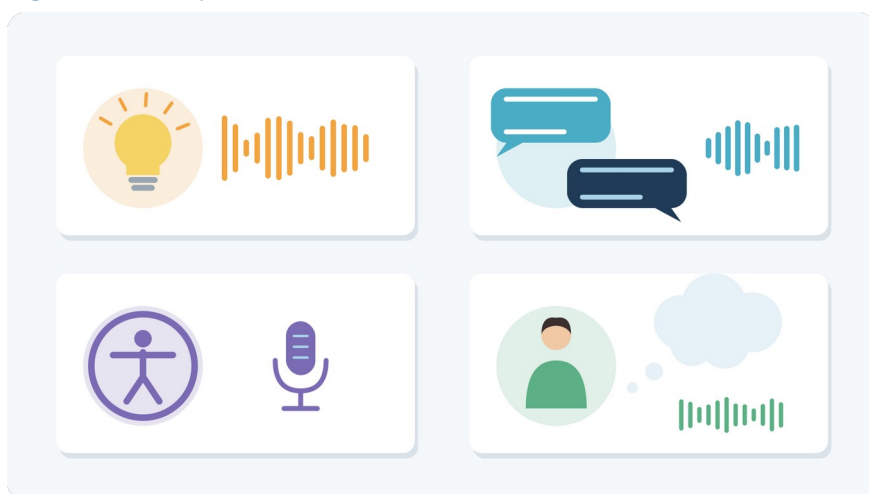
A hangmodul gyorsan fejlődött. 2025 vége óta a hangalapú beszélgetés közvetlenül a csevegésen belül zajlik, így a szóbeli párbeszéd és a vizuális tartalmak (például térkép, táblázat vagy kép) ugyanabban a beszélgetésben jelennek meg. 2026 júliusában az OpenAI bemutatta a **GPT-Live** hangmodell-generációt, amely egyszerre képes hallgatni és beszélni. Ettől természetessé válik a beszélőváltás és a közbevágás: beleszólhatunk, a modell jelezheti, hogy figyel, és a beszélgetés inkább egy emberi párbeszédre hasonlít.



Kétirányú (full-duplex) hangalapú beszélgetés: a felhasználó és az MI egyszerre beszélhet és hallgathat

- Az **élő mód** (Live, a legtöbb felhasználónál alapértelmezett) természetes oda-vissza beszélgetést tesz lehetővé, használhat internetes keresést és memóriát; összetett gondolkodást vagy kutatást igénylő feladatnál átadhatja a munkát egy erősebb szöveges modellnek, miközben a beszélgetés folytatódik.
- A **fejlett mód** (Advanced) megőrzi a korábbi valós idejű képességeket, köztük a kamera vagy a képernyő megosztását mobil eszközön – ez hasznos, ha egy tárgyat, eszközt vagy problémát szeretnénk megmutatni.
- A hangmodul **Projekteken** belül is használható, így a szóbeli beszélgetés a projekt fájljaira és utasításaira támaszkodhat.
- A beszélgetés után **átirat** jelenik meg a beszélgetési előzményekben. Ez nem feltétlenül egyezik pontosan az elhangzottakkal, főleg egymásba vágó beszéd, háttérzaj vagy gyors tempó esetén.
- A munkamenetekre az előfizetéstől függő használati korlátok vonatkoznak. A beszédfelismerés nyelve a beállításokban megadható, ami javítja a pontosságot.

Lehetséges szerepek a tanításban és a tanulásban



Ötletelés, nyelvgyakorlás, akadálymentesség és reflektív párbeszéd – a hangalapú interakció négy tipikus felhasználása

Ötletelés és tervezés útközben

Sok tanárnak ingázás vagy séta közben jutnak eszébe a legjobb ötletek. A szóbeli beszélgetés gyors módja annak, hogy hangosan gondolkodjunk egy tanóráról: „Holnap elsőéves tanulóknak tanítom a kerékcserét. Segíts motiváló órakezddést találni.” Utána kérjünk írásos összefoglalót a beszélgetésben, és szerkesszük át.

Nyelvgyakorlás

A hangmodul különösen értékes a szakmai kommunikáció idegen nyelvű gyakorlásához: a recepciós köszönti a szállodai vendéget, a szerelő elmagyarázza az ügyfélnek a javítást, a pincér felveszi a rendeltetést. Az MI eljátszhatja a partner szerepét, beszélhet lassabban vagy gyorsabban, ismételhet, és visszajelzést adhat a szóhasználatról. A kiejtéssel kapcsolatos visszajelzést fenntartással kezeljük; a viszonyítási pont a tanár marad.

Akadálymentesség

A hangalapú interakció csökkentheti az akadályokat az olvasási nehézséggel, látássérüléssel vagy mozgáskorlátozottsággal élő tanulók számára: olvasás és gépelés nélkül kérdezhetnek, és kaphatnak magyarázatot. Számukra a hangmodul valódi támogatás lehet – az iskola segítő technológiákra vonatkozó keretein belül.

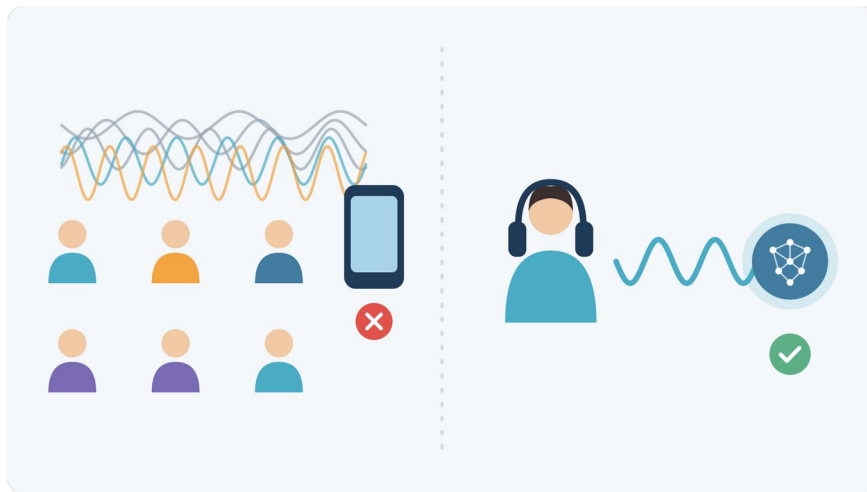
Gyors szakmai segítség

A tanműhelyben és a gyakorlati helyszíneken gyakran foglalt a kezünk. Egy szóban feltett kérdés („Általában mekkora meghúzási nyomatékot írnak elő...?”) gyorsabb lehet a gépelésnél. A biztonság szempontjából kritikus információt azonban mindig ellenőrizni kell a gyártói dokumentációban vagy a vonatkozó szabványban – egy magabiztos szóbeli válasz sem megbízhatóbb egy írotttnál.

Reflektív párbeszéd

A hangalapú beszélgetés jól alkalmas a reflexióra: egy nem terv szerint alakult tanóra átbeszélésére, egy nehéz szülői beszélgetésre való felkészülésre, vagy – a tanulók esetében – állásinterjú gyakorlására. Ha arra kérjük az MI-t, hogy kérdező „kritikus barátként” viselkedjen, ne pedig egyszerűen helyeseljen, a reflexió hasznosabb lesz.

Gyakorlati korlátok



Zajos tanteremben romlik a felismerés; a hangmodul csendes környezetben, négyszemközt működik a legjobban

- **Átírási hibák:** az akcentus, a tájszólás, a nyelvek keverése és a szakszavak félreérthetők. Ha fontosak a részletek, ellenőrizzük az átíratot.
- **Zaj:** a háttérzaj és az egyszerre beszélő emberek nem szándékolt válaszokat válthatnak ki. A hangmodul négyszemközt működik a legjobban; órán használjunk fejhallgatót vagy csendes sarkot, iPhone-on a hangelkülönítés (voice isolation) beállítás is segíthet.
- **Nem látszanak a források:** a szóbeli válasz nem mutatja olyan egyértelműen a forrásait, mint az írásos. Tényeknél kérjük el a forrásokat a beszélgetésben, és ellenőrizzük őket.
- **Meggyőző hangzás:** a gördülékeny, barátságos hang a hibás válaszokat is meggyőzőbbé teheti. A hallucináció hangalapú módban sem tűnik el.
- **Korlátok:** a munkamenet hosszára és a használatra korlátok vonatkoznak, egyes funkciók (pl. képernyőmegosztás) pedig csak bizonyos módokban érhetők el.

Adatvédelem és megfelelő használat a tanulókkal



A hangfelvételek és az átíratok személyes adatok: tájékozódjunk arról, hogyan tárolják őket, és ki férhet hozzájuk

- A ChatGPT-ben a hangalapú beszélgetések hangrészleteit az átirattal együtt tárolják a beszélgetési előzményekben, a szolgáltató szerint 30 napig; a modellek betanítására nem használják őket, hacsak a felhasználó nem dönt a megosztásuk mellett, vagy be nem kapcsolja a megfelelő beállítást. Használat előtt ellenőrizzük az aktuális feltételeket.
- Az ember hangja személyes adat. Ne rögzítsünk tanulókat vagy osztálytársakat hozzájárulás nélkül, és tartsuk be a GDPR-t és az iskola adatvédelmi szabályait.
- Tartsuk tiszteltben a szolgáltatók életkori feltételeit (a ChatGPT esetében legalább 13 év, 18 év alatt szülői hozzájárulással) és az iskola szabályait. Fiatalabb tanulóknál tanári bemutatót alkalmazzunk.
- Állapodjunk meg a szabályokban: a hangalapú beszélgetésekben ne legyen személyes vagy érzékeny információ, és minden használatnak legyen világos célja.

Tanulságok tanároknak

- Különböztessük meg a diktálást (beszédből szöveg) a hangalapú beszélgetéstől (szóbeli párbeszéd).
- A hangmodul erős eszköz az ötleteléshez, a nyelvgyakorláshoz, az akadálymentességhez és a reflexióhoz.
- Állítsuk be a nyelvet, használjuk csendes környezetben, és ellenőrizzük az átiratot, ha fontosak a részletek.
- A szóbeli válasz nem megbízhatóbb az írásosnál: ellenőrizzük a tényeket és a biztonsági információkat.
- Kezeljük a hangot személyes adatként: érvényesek a hozzájárulásra, az életkorra és az iskolai szabályokra vonatkozó előírások.

Hivatkozások

- OpenAI (2026): Introducing GPT-Live. OpenAI blog, 2026. július 8.
- OpenAI Help Center (2026): ChatGPT Voice; ChatGPT release notes.
- Lengyel Zs. (2026): Konkrét alkalmazások tartalomgenerálásra. VITA képzési prezentáció.

Linkek

- ChatGPT Voice – GYIK: help.openai.com/en/articles/8400625-voice-mode-faq
- A GPT-Live bemutatása: openai.com/index/introducing-gpt-live

Szerzői jog © 2026 VITA Konzorcium

Ez a mű a Creative Commons Nevezd meg! – Így add tovább! 4.0 Nemzetközi Licenc (CC BY-SA 4.0) feltételei szerint újra felhasználható és átdolgozható.





Az Európai Unió finanszírozásával. A kifejtett nézetek és vélemények kizárólag a szerző(k)éi, és nem feltétlenül tükrözik az Európai Unió vagy a Nemzeti Ügynökség álláspontját.

