Cacciato, A. & Looi, J. (Accepted). Listening and data-driven learning. In L. McCallum & D. Tafazoli (Eds.), *The Palgrave Encyclopedia of Computer-Assisted Language Learning*. Springer.

Listening and Data-Driven Learning

Abstract

Listening, the most used language skill in daily life, remains underrepresented in both instructed second language acquisition (ISLA) and data-driven learning (DDL) research. This entry explores how DDL can enhance listening instruction by granting learners access to spoken or multimodal corpora, enabling engagement with authentic language data. It distinguishes listening as oral comprehension from auditory-phonetic processing, examines challenges such as natural speech features, segmentation difficulties, and instructional limitations, and discusses strategies to facilitate listening instruction with corpora. Additionally, it highlights how multimodal corpora and tools like YouGlish and the TED Corpus Search Engine support listening development through pattern recognition and perceptual training. These corpora feature authentic language data, helping learners process connected speech, recognise discourse functions, and refine pronunciation, making learning more relevant and motivating. The discussion emphasises the need for further research and corpus development, particularly for languages other than English, and for a more balanced integration of listening in DDL and L2 pedagogy.

Keywords Listening, Pronunciation, Oral comprehension, Data-driven learning, Corpora, Multimodal learning

1 Introduction

Listening, despite being the most frequently used language skill in daily life (Morley, 2001), has received comparatively less attention than speaking in both instructed second language acquisition (ISLA) (McCarthy & McCarten, 2023) and data-driven learning (DDL) research, where it remains the least explored among the four core language skills (Boulton & Vyatkina, 2021). Yet, listening plays a crucial role in language learning, serving as the primary means through which learners receive linguistic input. As the process of listening comprehension requires the integration of phonological, lexical, syntactic, semantic, and pragmatic competences (Flowerdew, 1994), L2 learners often encounter challenges.

Research on listening has emphasised the importance of exposure to a large amount of authentic speech to develop real-world listening skills (Camiciottoli, 2023). DDL, granting learners access to spoken and written language samples, presents an innovative way to enhance listening skills. By interacting with authentic language data, learners analyse concordances, linguistic patterns, and frequency distributions to gain insights into discourse genres and socio-cultural values (Wang & Wu, 2014). This process helps develop listening skills while fostering awareness of grammatical structures and discourse patterns (Arslanbay & Ceylan, 2023).

2. L2 listening instruction

2.1 Listening as oral comprehension and auditory-phonetic processing

Evaluating the potential of DDL for teaching listening may begin with distinguishing between ‘oral comprehension’, focusing on deciphering information in connected speech, and ‘auditory-phonetic processing’, relating to perceiving and producing accurate pronunciation and prosody.

Developing auditory-phonetic processing requires perceptual training, which enhances not only the perception of segmental features (i.e., consonants and vowels) and suprasegmental features (i.e. lexical stress, pitch accent, rhythm, and intonation) but also their production (Lengeris, 2012). Mastering these features is essential for achieving clear and natural-sounding speech (Ma et al., 2024), with intelligibility emerging as the primary focus of pronunciation teaching research over the past fifteen years (Qian & Deris, 2023).

In contrast, listening as oral comprehension involves identifying meaning in connected speech and its communicative functions (Cuq & Gruca, 2005), drawing on linguistic competencies (Flowerdew, 1994) and cultural knowledge (Vandergrift, 2007). Listeners process auditory input through top-down (accessing contextual information to construct meaning) and bottom-up (analysing smaller lexical and syntactic units) approaches, with comprehension emerging from the interaction of both (Stæhr, 2009). As communication is multimodal (O'Halloran, 2011), listening processes both auditory and visual input, incorporating gestures, facial expressions, gaze, and body posture (Jewitt, 2013). Oral comprehension follows a three-phase cognitive process: perceptual process (encoding auditory and visual input), parsing (constructing a mental representation of meaning), and utilisation (integrating this representation into long-term memory) (Anderson, 1995).

2.2 Issues in listening instruction

The cognitive demands of listening pose significant challenges for both L2 learners and instructors. As Reed & Liu (2020) point out, learners face difficulties in parsing continuous speech and identifying speaker intent, with issues such as understanding rapid, fluent speech and recognising non-prominent or truncated words that deviate from their citation forms. A major obstacle is the features of natural speech, such as elision, assimilation, and weak forms, which can hinder comprehension. Furthermore, the lack of clear pauses in natural speech makes segmenting words difficult (Madarbakus-Ring, 2020). For instructors, a key challenge remains the tendency to focus solely on assessing comprehension accuracy rather than actively teaching listening skills (Field, 2008).

Research highlights a shift in oral skills research towards production, with listening comprehension being sidelined (Siegel, 2015). Teachers often overestimate the difficulty of authentic listening materials, even though factors like background noise, digressions, and syntactic ruptures are not major barriers (Janin, 2019). However, pace and rhythm present significant challenges. Moreover, any classroom listening techniques fail to adequately prepare learners to handle the complexities of natural speech.

Given the close connection between speaking and listening, pronunciation plays a crucial role in comprehension. Clear pronunciation supports the processing of fluent, connected speech and benefits other language skills such as reading and writing. As Ma et al. (2024) highlight, mispronunciations can lead to misunderstandings, whereas intelligible pronunciation facilitates communication.

3. DDL, corpora, and listening instruction

3.1 Multimodality of listening

Multimodal or spoken corpora reveal key features of spoken language, such as ellipsis, hesitation, overlaps, discourse markers, and idiomatic expressions; engaging with these features in context help learners recognise and process them more effectively, improving comprehension in turn (Campbell et al., 2007). These corpora show authentic dialogues, with features like dialogic turns, conversation openings, and colloquial responses, offering insights into pragmatic phenomena such as courtesy markers while revealing grammatical structures, discourse patterns and prosodic information, making them a more authentic representation of language use than written texts. Furthermore, exposing the learner to learning material that closely mirrors reality will make the learning process more relevant to them, motivating them to improve their target language skills (Campbell et al., 2007).

Multimodal corpora integrate audio, video, and textual data, allowing for the analysis of both linguistic and non-linguistic elements, and offer learners visual and verbal cues that reinforce language learning (Coccetta, 2023). The emergence of DDL has broadened the scope of

multimodal learning (Aston, 2015). DDL with multimodal corpora combines textual, auditory, and visual elements to improve both speaking and listening skills, providing learners with an immersive experience in authentic language use (Boulton & Cobb, 2017). This approach not only fosters learner autonomy but also makes language learning more engaging and contextually relevant.

Research on multimodal corpora provides insights into how these modes function together, forming the foundation for innovative approaches to listening instruction. Multimodal listening entails both auditory and visual engagement, with meaning emerging from the interplay of different semiotic elements (Campoy-Cubillo & Querol-Julián, 2015) video content provides exposure to authentic language in real-life scenarios, familiarising learners with various accents, colloquial expressions, and situational language use (Sari & Margana, 2019). Kim (2016) highlights that visually dynamic and engaging video content captures learners' attention, increases motivation and enhances learning experience. Additionally, it facilitates a deeper understanding of cultural subtleties, social interactions, and non-verbal communication, leading to a more comprehensive grasp of the target language (Eom, 2024).

3.2 Corpus Resources for Listening

Innovative multimodal corpora and online resources have been developed to offer audio-visual resources to analyse listening features in context, such as:

- YouGlish (https://youglish.com) is an online learning tool that uses real-world YouTube videos to provide pronunciation examples in up to 21 languages. It offers features like adjustable playback speed, phonetic transcriptions, and personalised learning tracks, tailored for pronunciation, vocabulary, and comprehension instruction.

- The TED Corpus Search Engine (TCSE https://yohasebe.com/tcse) (Hasebe, 2015) is a valuable resource for exploring multimodal aspects of language, featuring transcripts and translations of over 1,800 TED Talks. It allows users to query the text, providing access to videos aligned with transcripts. The TCSE supports searches by surface text forms, lemmas, and parts of speech, making it an essential tool for academic research and language instruction.

- TED Talks (https://www.ted.com/talks), are a series of presentations that cover a wide range of topics, disseminating insights and ideas from experts in various fields. These talks are valuable resources for improving academic listening skills, expanding vocabulary, and understanding complex concepts. With their authenticity and diversity, TED Talks offers learner-friendly format for engaging with academic content and a rich source of academic spoken vocabulary.

A non-exhaustive overview of other freely available resources for L2 listening development is provided in Table 1, focusing on English language learners or on learners of different target languages.

Corpus	Type	Size	Description/Focus	Access
British National Corpus (BNC)	General	100 million	British English from the 1990s, various registers (written and spoken)	http://www.natcorp.ox.ac.uk

Trinity-Lancaster Corpus (TLC)	Specialised	4.2 million	Interactions between English speakers and learners during oral exams	https://cass.lancs.ac.uk/trinity-lancaster-corpus/
French Learner Language Oral Corpora (FLLOC)	Specialised	4000 files	Sub-corpora of French learners, includes audio files and tagged transcripts	https://web-archive.southampton.ac.uk/www.flloc.soton.ac.uk/search.php.html
LANGSNAP	Specialised	1141 recordings	Semi-structured interviews and story retelling tasks by learners of French and Spanish	https://web-archive.southampton.ac.uk/langsnap.soton.ac.uk/index.html
Spanish Learner Language Oral Corpora (SPLLOC)	Specialised	461 learners and 100 English speakers recordings	Includes narratives, interviews, and pedagogical tasks by English-speaking learners of Spanish	https://web-archive.southampton.ac.uk/www.splloc.soton.ac.uk/
European Science Foundation Second Language SLA Bank	General	Variable	Unplanned speech from adult immigrant workers interacting with speakers of various languages	https://slabank.talkbank.org/access/Multiple/ESF/
MICASE (Michigan Corpus of Spoken Academic English)	Specialised	1.8 million	Academic discourse from lectures, seminars, and advising sessions	https://quod.lib.umich.edu/cgi/c/corpus/corpus?c=micase;page=mbrowse
BASE (British Academic Spoken English Corpus)	Specialised	1.75 million	Recordings of lectures and seminars at UK universities	https://ota.bodleian.ox.ac.uk/repository/xmlui/handle/20.500.12024/2525
FLEURON	Specialised	Variable	Resource for foreign learners planning to study in France, with multimedia tools	https://fleuron.atilf.fr
FLORALE	Specialised	Variable	Orthographic transcriptions of Swiss and French radio programs.	https://florale.unil.ch

Table 1 List of existing resources for DDL listening instruction

3.3 What corpora can offer in listening instruction

Although communication involves at least two interlocutors, the focus in ISLA has historically prioritised speakers and speaking skills, often treating oral production as the product of an individual speaker rather than a speaker-listener dynamic (McCarthy & McCarten, 2023). However, the listener plays an active role in communication, establishing the impact of the

speaker's speech (Munro & Derwing, 2015). Corpus analysis shows that the listener contributes without taking the floor through non-verbal behaviours (e.g., head nods, hand positions), vocal but non-verbal responses (e.g., *uh-huh*, *aha*), single verbal items (e.g., *yeah*, *right*), complex affirmations (e.g., *that's true*), and pseudo-interrogatives (e.g., *Oh really?*), highlighting the need to integrate 'listenership' into speaking and listening pedagogy, acknowledging the speaker-listener partnership (McCarthy & McCarten, 2023) or the listener's role in 'bi-directional listening' (Vandergrift, 2007). Additionally, a listener's background – such as accent familiarity, content knowledge, attitudes, language proficiency, and linguistic awareness – affects communication and intelligibility (Ginther & Yan, 2017). Teaching listening should therefore emphasise mutual intelligibility, ensuring learners can both make themselves understood and comprehend others (Levis, 2018).

3.3.1 Authenticity and naturalness

Research on listening comprehension emphasises linguistic authenticity and representativeness, as the goal of instruction is to prepare learners for real-world communication (Vandergrift, 2007). Language serves two functions: transactional for communicating information, and interactional for expressing social relations and attitudes (Brown & Yule, 1983). Transactional language is typically planned, while interactional language is more spontaneous (Mordaunt & Olson, 2010). Mordaunt & Olson (2010) propose a 'comprehension corpus' incorporating both transactional and interactional speech, planned and unplanned discourse, and diverse cultural themes (e.g., family, emotions, work, shopping) across various genres and styles (e.g., news broadcasts, talk shows, variety shows). Authentic corpora expose learners to natural speech features (Campbell et al., 2007). This exposure enhances naturalness in conversation, not as a pursuit of natural-like performance (Warren, 2006), but as alignment with expected social norms (McCarthy & McCarten, 2023).

Corpus linguistics and data-driven learning further aid learners by identifying overused or underused items and elements that contribute to foreign-soundingness even in the absence of grammatical errors (Granger, 2004). Exposure to authentic speech enhances metalinguistic awareness (e.g., contraction, linking, sound reduction, stress, pitch, intensity) and extralinguistic awareness, such as how intonation conveys speaker attitudes and emotions (Reed & Liu, 2020).

3.3.2 Lexical coverage and vocabulary profiling

Processing auditory input is an inferential process where the listener draws on linguistic, paralinguistic, and cultural knowledge (Vandergrift, 2007). Lower-proficiency listeners rely more on bottom-up processing, focusing on smaller linguistic units (e.g., lexical syntactic), while more skilled listeners use cognitive and metacognitive strategies, leveraging top-level cues (e.g., contextual information) but also depending on a strong lexical foundation (Stæhr, 2009). Thus, effective comprehension requires lexical coverage – 95% coverage of the 4,000-5,000 word families for adequate understanding and 98% of the 8,000-9,000 words families for high-level comprehension (Nurmukhamedov, 2017).

Therefore, effective listening instruction requires selecting materials at an appropriate lexical level. Online vocabulary profiling tools (e.g., LexTutor) can help grade text difficulty based on word frequency, word families, and word lists (Madarbakus-Ring, 2020). Romanelli et al.'s (2014) comparison of TED Talks and academic lectures offers insights into selecting authentic, level-appropriate resources. Additionally, unfamiliar proper nouns hinder comprehension (Kobeleva, 2012), and pre-teaching them using word lists can improve learners' oral comprehension (Nurmukhamedov, 2017).

4 DDL strategies for listening

Processing aural input in real time is challenging for L2 learners due to working memory limitations (Stæhr, 2009). The post-listening phase is crucial because it allows learners to engage with the material beyond answering comprehension questions (Campbell et al., 2007), focusing on listening as a process (Field, 2003). Multimodal corpora that enable audio replay and slow playback helps learners identify difficult segments and notice features of fast speech (e.g., elision, assimilation, stress) (Campbell et al., 2007).

4.1 Critical listening, shadowing and mirroring

Perceptual loop theory (Levelt, 1983) suggests that learners continuously compare their speech to aural input, forming a feedback loop. They also use their own auditory feedback to refine L2 phonetic categories, which enhances perceptual ability (Linebaugh & Roche, 2015). A key strategy for self-monitoring is 'critical listening', where learners compare incorrect and correct pronunciation in audio input (Fraser, 2001). Recording and relistening to their own speech aids analysis and refinement. Additional techniques include shadowing (repeating speech in real time) and mirroring (mimicking a speaker's gestures, expressions, and speech simultaneously) (Qian & Deris, 2023).

4.2 Chunks and chunking

Language is constructed and learned in chunks – units larger than individual words (Lightbown & Spada, 2006). In listening, learners group aural input into semantic chunks, while advanced listeners process larger segments of information (Mordaunt & Olson, 2010). Chunking correlates with speaking rate: at faster speeds, chunks like *that's right* serve as conversational feedback, while at slower speeds, they take on a more pragmatic role, signalling personal and active engagement (Campbell et al., 2007). Recognising chunk functions enhances information retrieval and helps learners focus on the broader discourse structure (Nattinger & DeCarrico, 1992).

Pattern hunting (cf. Kennedy & Miceli, 2017) is an effective listening strategy where learners identify phraseological items, examine their collocations, semantics, pragmatics, and prosody, and compare them to L1 equivalents (Friginal et al., 2020). This approach is particularly useful for academic listening, helping learners recognise signalling nouns (Flowerdew, 2003), linguistic and extra-linguistic features (Camiciottoli, 2007), culture-specific references and informal expressions (Camiciottoli, 2020), metadiscourse (Molino, 2018), keywords, multiword units, and patterns for interaction (Janin, 2019) for understanding the organisation of academic talks, interactional purposes such as engaging with participants and referencing multimodal resources. Using multimodal resources like videos with captions help unpack chunks, activate selective and global listening strategies, and enhance processing automaticity (Winke et al., 2010).

A structured approach benefits listening instruction (Hamada, 2019). Established DDL frameworks, such as Illustration-Interaction-Intervention-Induction (Carter & McCarthy, 1995; Flowerdew, 2009), support chunk learning by guiding learners to extract, interact with, and incorporate representative conversational patterns into their language repertoire (McCarthy & McCarten, 2023). Teachers intervene as needed to address difficulties, ensuring effective learning.

5.0 DDL task recommendations for listening

Grounded in the distinction between listening as auditory-phonetic processing and listening as oral comprehension, this section illustrates how existing corpora can assist in listening instruction through DDL.

5.1 Pattern hunting for listening comprehension

Using TCSE, learners can identify and analyse the role of discourse markers and metadiscourse markers, which serve distinct functions in academic speech. Discourse markers (e.g., *however*, *in contrast*, *therefore*) primarily signal logical connections, topic shifts, and discourse organisation, while metadiscourse markers (e.g., as I mentioned before, what I mean is, *let's move on to*) help structure speech, guide the listener, and enhance engagement. Examining these markers in context allows learners to explore their collocations, meanings, pragmatic functions, and prosodic features, reinforcing their understanding of how chunks contribute to textual coherence.

A systematic approach to integrating TCSE into classroom activities or independent practice involves the following steps:

(1) Select a short video from the TCSE corpus using the *List all talks* function, ensuring the speech speed and readability match the learner's level;

(2) Listen to the talk without the transcript, focusing on overall comprehension, and answer questions to identify the key points of the talk;

(3) Activate the *Discourse markers* function, which automatically detects discourse markers and metadiscourse markers within the transcript;

(4) Analyse these markers by examining their placement within the sentence, and their functions (e.g. idea introduction, logical transitions);

(5) Use the *Talks of highest cosine similarity* function to compare different examples of the same pattern, noting recurrent expressions and their pragmatic roles in speech;

(6) Apply the pattern hunting approach to identify keywords and metadiscourse markers, comparing their usage with equivalent structures in the learner's L1;

(7) Use the *Show talk keywords* and *Talks of highest cosine similarity* functions to extract relevant linguistic patterns and observe their distribution across different contexts;

(8) Leverage the *Construction* function to analyse metadiscourse markers within chunks, considering their common collocations, syntactic structure, length, and prosodic variations;

(9) Focus on identifying occurring 2- to 4-word sequences (n-grams) recognising structural patterns around key phrases and exploring variations in discourse marker placement;

(10) Listen to a new lecture, different from the one analysed initially, and answer comprehension questions to reinforce their understanding.

However, no single tool can fully support the study of spoken discourse in languages other than English (LOTEs), for which a combination of multiple resources is often required. For instance, when working with French, learners can use a pre-selected YouTube video alongside FLORALE, a corpus allowing for examining conversation markers while taking into account sociocultural factors. A similar protocol as listed can be established, focusing on one key feature of oral discourse. Additionally, YouGlish (Section 5.2) can serve as a concordancer for retrieving and analysing specific instances of discourse markers in naturally occurring spoken French, helping learners compare their usage across various contexts. By integrating multiple tools, learners gain a more comprehensive understanding of discourse structures in different languages, beyond what any single existing corpus or automated function can provide.

5.2 Post-listening perceptual training

In the post-listening phase, where learners focus on listening as a process, YouGlish is a valuable tool. Though not a traditional corpus, it provides authentic YouTube videos based on

search terms. Videos are preloaded a few seconds before the term appears, with the term highlighted in KWIC format, and the full transcript accessible via the caption window.

YouGlish supports 21 languages (e.g., Chinese, French, German, Greek, Japanese, Polish, Thai) and allows switching between regional varieties (e.g., French in France, Canada, and Belgium; English in the US, UK, and Australia; Spanish in Spain and Latin America). Unlike traditional dictionaries, it enables searches for phrases (e.g., *not at all*) and intonation patterns using a question mark or exclamation mark (e.g. *Absolutely!*). For each search term, YouGlish provides nearby words, IPA transcriptions, and similar pronunciations under the video for further exploration.

Creating a free personal account unlocks additional learner-focused features, such as saving past searches, bookmarking useful videos, and adding words, expressions, notes, and definitions in 'My content'. Learners can also record themselves using the *record* function, with YouGlish storing up to three recordings for comparison with the search term in the video.

These functions facilitate post-listening perceptual training. Learners can search for words and listen repeatedly to identify segmental features – e.g., how *idea* [ʌɪˈdɪə] is pronounced with distinct vowels and in three syllables, how *woman*[ˈwʊmən] and *women* [ˈwɪmɪn] differ in their first syllable, how /d/ is dropped in *handsome* [ˈhans(ə)m], and how *fast* [fɑːst] is reduced in *breakfast* [ˈbrɛkfəst]. These segmental features, shaped by stress shifts, historical sound change, and vowel reduction, highlight the complexity of spoken language and the usefulness of tools like YouGlish.

YouGlish's ability to search for phrases makes it a powerful tool for perceptual training, allowing learners to observe suprasegmental features like elision, assimilation, and weak forms beyond citation forms. Through repeated exposure, learners can identify, for example, the elision of *and* reducing to a schwa in fish and chips [ˈfɪʃ ən tʃɪps]; the weakening of *to* in *what to do* [wɒt tə duː]; the assimilation of /t/ and /j/ into /dʒ/ in *would you like* [wʊdʒu laɪk]; or the coarticulation and flapping in *not at all* by comparing British [nɒt ət ɔːl] and American English [nɑːɾəˈtɔːl].

YouGlish can be systematically integrated into classroom activities following oral comprehension tasks or encouraged as independent practice when learners acquire new vocabulary. Below is a widely applicable framework for implementation:

(1) Identify words with challenging pronunciation from a comprehension task or newly encountered vocabulary;

(2) Search on YouGlish to examine segmental features (consonants and vowels);

(3) Expand the search to nearby words or chunks to observe suprasegmental features (e.g., stress, elision, assimilation, linking);

(4) Compare pronunciations across different regional varieties;

(5) Analyse patterns and formulate pronunciation rules based on observations;

(6) Practice shadowing by repeating words simultaneously or immediately after the video;

(7) Use the *record* function to check pronunciation accuracy;

(8) (Optional) Engage in mirroring by imitating gestures, facial expressions, and body language.

6 Conclusion

Listening is the most frequently used language skill in daily life (Morley, 2001) and is essential for language learning. However, the division of language into four discrete skills (i.e. reading,

writing, speaking, and listening) has led to listening being side-lined, particularly in comparison to speaking. This trend is reflected in DDL research, where Boulton and Vyatkina (2021) found that among 125 studies explicitly targeting one of the four skills, only four focused on listening over the past three decades.

The limited focus on listening is both a cause and a consequence of the underdevelopment of corpora and resources designed for listening instruction. However, as this entry highlights, existing research and resources can support a stronger integration of DDL into listening pedagogy. Going forward, it is essential to develop more corpora and resources, particularly for LOTEs, and to distinguish between oral comprehension and auditory-phonetic processing in listening instruction. Further research at the intersection of DDL and listening is needed to promote uptake and encourage a more balanced development of all four language skills in L2 pedagogy.

References

Anderson, J.R. (1995). *Cognitive Psychology and its Implications* (4th ed.). Freeman.

Arslanbay, G., and Ersanli, C. Y. (2023). It's a Long Story, but DDL is worth it. Data-Driven Learning as Multimodel Method for English Sessions in Turkish Prep-Classes. *i-manager's Journal on English Language Teaching*, 13(2), 1-15. https://doi.org/10.26634/jelt.13.2.19343

Aston, G. (2015). Learning phraseology from speech corpora. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 65–84). Benjamins.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393.

Boulton, A. & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning & Technology*, 25(3), 66-89.

Brown, G. & Yule, G. (1983). *Discourse Analysis*. Cambridge University Press.

Camiciottoli, B.C. (2023). Corpora and EAP listening comprehension. In R. Jablonkai & E. Csomay (Eds.), *The Routledge Handbook of Corpora and English Language Teaching and Learning*, (pp. 248-263). Routledge. https://doi.org/10.4324/9781003002901

Camiciottoli, C.B. (2007). *The Language of Business Studies Lectures: A Corpus-assisted Analysis*. Benjamins.

Camiciottoli, C.B. (2020). The OpenCourseWare lecture: A new twist on an old genre? *Journal of English for Academic Purposes*, 46, 100870. https://doi.org/10.1016/j.jeap.2020.100870

Campbell, D.F., McDonnell, C., Meinardi, M. & Richardson, B. (2007). The need for a speech corpus. *ReCALL*, 19(1), 3-20. https://doi.org/10.1017/S0958344007000213

Campoy-Cubillo, M.C. & Querol-Julián, M. (2015). Assessing multimodal listening. In B. Crawford & I. Fortanet-Gómezn (Eds.). *Multimodal Analysis in Academic Settings: From Research to Teaching*. London/ New York: Routledge, 193-212.

Carter, R. & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, 16(2), 141-158. https://doi.org/10.1093/applin/16.2.141

Coccetta, F. (2023). Multimodal corpora and concordancing in DDL. In R. Jablonkai & E. Csomay (Eds.), *The Routledge Handbook of Corpora and English Language Teaching and Learning*, (pp. 248-263). Routledge. https://doi.org/10.4324/9781003002901

Cuq, J.P. & Gruca, L. (2005). *Cours de Didactique du Français Langue Étrangère et Seconde*. CLE International.

Eom, M. (2024). Empowering English-Speaking Skills through Personalized Exploration and

Application of 'YouGlish.' *Korean Journal of English Language and Linguistics*, 24, 236–256. https://doi.org/10.15738/kjell.24..202403.236

Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal*, 57(4), 325-333. https://doi.org/10.1093/elt/57.4.325

Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.

Flowerdew, J. (1994). Research of relevance to L2 lecture comprehension. In J. Flowerdew (Ed.), *Academic Listening. Research Perspectives*, (pp. 7-29). Cambridge University Press.

Flowerdew, J. (2003). Signalling nouns in discourse. *English for Specific Purposes*, 22(4), 329-346.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393-417. https://doi.org/10.1075/ijcl.14.3.05flo

Fraser, H. (2001). *Teaching Pronunciation: A Handbook for Teachers and Trainers*. *Three Frameworks for an Integrated Approach*. DETYA.

Friginal, E., Dye, P. & Nolen, M. (2020). Corpus-based approaches in language teaching: Outcomes, observations, and teacher perspectives. *Bogazici University Journal of Education*, 37(1), 43-68.

Ginther, A. & Yan, X. (2017). Listeners and raters: Similarities and differences in evaluation of accented speech. In O. Kang & A. Ginther (Eds.), *Assessment in Second Language Pronunciation*, (pp. 79-100). Routledge.

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In U. Connor & T. Upton (Eds.), *Applied Corpus Linguistics: A Multidimensional Perspective*, (pp. 123-145). Rodopi.

Hamada, Y. (2019). Shadowing: What is it? How to use it. Where will it go?. *RELC Journal*, 50(3), 386-393. https://doi.org/10.1177/0033688218771380

Hasebe, Y. (2015). Design and implementation of an online corpus of presentation transcripts of TED Talks. *Procedia – Social and Behavioral Sciences*, 198, 174–182.

Janin, L. (2019). La comprehension orale des conferences universitaires : Approche linguistique et didactique. *Travaux Neuchâtelois de Linguistique*, 71, 45-65.

Jewitt, C. (2013). Multimodal teaching and learning. In C.A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*, (pp. 4109-4114). Blackwell.

Kennedy, C. & Miceli, T. (2017). Cultivating effective corpus use by language learners. *Computer Assisted Language Learning*, 30(1-2), 91-114. https://doi.org/10.1080/09588221.2016.1264427

Kim, Y. S. G. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology*, 141, 101-120. https://doi.org/10.1016/j.jecp.2015.08.003

Kobeleva, P.P. (2012). Second language listening and unfamiliar proper names: Comprehension barrier?. *RELC Journal*, 43(1), 83-98. https://doi.org/10.1177/0033688212440637

Lengeris, A. (2012). Prosody and second language teaching: Lessons from L2 speech perception and production research. In J. Romero-Trillo (Ed.), *Pragmatics and Prosody in English Language Teaching*, (pp. 25-49). Springer.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41-104.

Levis, J.M. (2018). *Intelligibility, Oral Communication, and the Teaching of Pronunciation*. Cambridge University Press.

Lightbown, P.M. & Spada, N. (2006). *How Languages are Learned*. Oxford University Press.

Linebaugh, G. & Roche, T. (2015). Evidence that L2 production training can enhance perception. *Journal of Academic Language & Learning*, 9(1), A1-A17.

Ma, Q, Mei, F. & Qian, B. (2024). Exploring EFL students' pronunciation learning supported by corpus-based language pedagogy. *Computer Assisted Language Learning*, 1-27. https://doi.org/10.1080/09588221.2024.2432965

Madarbakus-Ring, N. (2020). Developing graded TED Talks to integrate academic vocabulary into listening lessons for pre-sessional learners. In *The TESOL Encyclopedia of English Language Teaching*, (pp. 19-7). John Wily & Sons. https://doi.org/10.1002/9781118784235.eelt0990

McCarthy, M. & McCarten, J. (2023). Speaking and listening: Two sides of the same coin. In K. Harrington & P. Ronan (Eds.), *Demystifying Corpus Linguistics for English Language Teaching*, (pp. 59-78). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-11220-1

Molino, A. (2018). 'What I'm speaking is almost English…': A corpus-based study of metadiscourse in English-medium lectures at an Italian university. *Educational Sciences: Theory and Practice*, 18(4), 935-956.

Mordaunt, O. & Olson, D. (2010). Listen, listen and listen: Building a comprehension corpus and making it comprehensible. *Educational Studies*, 36(3), 249-258. https://doi.org/10.1080/03055690903172068

Morley, J. (2001). Aural comprehension instruction: Principles and practices. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language*, (pp. 69-85). Heinle and Heinle.

Munro, M.J. & Derwing, T.M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. Levis (Eds.), *The Handbook of English Pronunciation*, (pp. 377-369). John Wiley & Sons.

Nattinger, J.R. & DeCarrico, J.S. (1992). *Lexical Phrases and Language Teaching*. Oxford University Press.

Nurmukhamedov, U. (2017). Lexical coverage of TED Talks: Implications for vocabulary instruction. *TESOL Journal*, 8(4), 768-790. https://doi.org/10.1002/tesj.323

O'Halloran, K.L. (2011). Multimodal discourse analysis. In K. Hyland & B. Paltridge (Eds.), *The Bloomsbury Companion to Discourse Analysis*, (pp. 120-137). Continuum.

Qian, B. & Deris, F.D. (2023). A corpus-based pronunciation teaching model: A conceptual paper. *Arab World English Journal*, 14(1), 71-88. https://dx.doi.org/10.24093/awej/vol14no1.5

Romanelli, F., Cain, J., & McNamara, P. J. (2014). Should TED talks be teaching us something? *American Journal of Pharmaceutical Education, 78*(6), 113-116. https://doi.org/10.5688/ajpe786113

Reed, M. & Liu, D. (2020). Technology-enhanced L2 listening: Triangulating perception, production, and metalinguistic awareness. In O. Kang, S. Staples, K. Yaw & K. Hirschi (Eds.), *Proceedings of the 11th Pronunciation in Second Language Learning and Teaching Conference*, (pp. 173-185). Iowa State University.

Sari, Y. N., & Margana, M. (2019). YouTube as a learning media to improve the student's speaking ability in 21st century. *Journal of English Language Teaching and Linguistics*, *4*(2), 263.

Siegel, J. (2015). *Exploring listening strategy instruction through action research*. Palgrave

Macmillan. https://doi.org/10.1057/9781137521903.

Stæhr, L.S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(4), 577-607. https://doi.org/10.1017/S0272263109990039

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191-210.

Wang, Y. & Wu, L. (2014). College English classroom teaching design based on DDL. *2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control*, 109-122. https://doi.org/10.1109/imccc.2014.30

Warren, M. (2006). *Features of Naturalness in Conversation*. John Benjamins.

Winke, P., Gass, S. & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65-86. http://dx.doi.org/10125/44203