AL-AZHAR UNIVERSITY
BULLETIN OF THE FACULTY
OF
LANGUAGES & TRANSLATION

جامعة الأزهر
مجلة كلية اللغات والترجمة

Exploring the Impact of Technology on Perceptual Teaching of Greek Pronunciation: Comprehensive Review of CAPT Methods

By

D/ Youssef Yasser Abdelhakim Elsafoury

Department of Greek and Italian languages,
Faculty of Languages and Translation
Al-Azhar University, Cairo, Egypt

Exploring the Impact of Technology on Perceptual Teaching of Greek Pronunciation: Comprehensive Review of CAPT Methods

Youssef Yasser Abdelhakim Elsafoury
Department of Modern Greek and Italian Languages, Faculty of Languages and Translation, Al - Azhar University, Cairo, Egypt.
Email: youssefyasser.2110@azhar.edu.eg

ABSTRACT This research explores the influence of technology on perceptual teaching/training of Greek pronunciation, specifically focusing on Computer-Assisted Pronunciation Teaching/Training (CAPT) methods. The study reviews three prominent techniques: High Variability Phonetic Training (HVPT), audiovisual perceptual training using the SpeakGreek software, and Cued Pronunciation Reading (CPR). In the context of language learning, these methods can be implemented in various settings, including classrooms or laboratories. The central theme of each method revolves around enhancing learners' perceptual skills, thereby facilitating the production of both segmental elements and suprasegmental elements.

Keywords: Perceptual Training, CAPT, Cued Pronunciation Reading, HVTP, Greek Pronunciation.

تقصي تأثير استخدام التكنولوجيا في التدريس الادراكي للنطق اليوناني: نظرة شاملة على طرق تدريس النطق بمساعدة الحاسوب

يوسف ياسر عبد الحكيم الصافوري
قسم اللغات اليونانية والايطالية، كلية اللغات والترجمة، جامعة الازهر، القاهرة، مصر.
البريد الإلكتروني:youssefyasser.2110@azhar.edu.eg

الملخص: يتقصى هذا البحث تأثير التكنولوجيا على التعليم/التدريب لفهم المنطوق اليوناني، مع التركيز بشكل خاص على أساليب تعليم/تدريب النطق بمساعدة الكمبيوتر .(CAPT) تستعرض الدراسة ثلاث تقنيات مهمة: التدريب الصوتي المتباين(HVPT) ، والتدريب لفهم المنطوق من خلال المنهج السمعي البصري audiovisualباستخدام برنامج SpeakGreek، وسماع النصوص مع ارشادات بصرية تحدد التنغيم والوقف .(CPR). وفي سياق تعلم اللغة، يمكن تنفيذ هذه الطرق في بيئات مختلفة، بما في ذلك داخل الفصل، أو داخل المعمل، يدور الموضوع الرئيسي لكل طريقة حول تعزيز المهارات الإدراكية (فهم المنطوق) لدى المتعلمين، وبالتالي تسهيل إنتاج (التحدث) كل من العناصر القطعية والعناصر الفوقطعية

الكلمات المفتاحية: التدريب الإدراكي، تدريس النطق بمساعدة الحاسوب، القراءة النطقية مع ارشادات بصرية، التدريب الصوتي المتباين، نطق اللغة اليونانية.

1. Introduction

Language is a fundamental tool for communication and understanding the world around us. While acquiring our first language (L1) feels almost effortless, mastering the pronunciation of a second language (L2) or foreign language (FL) can be a herculean task. This stark difference highlights the fascinating complexity of human language learning. Children are naturally drawn to language. From the moment they are born, they are immersed in the sounds and rhythms of their surrounding environment, typically their mother tongue. According to research by Ohala (2008), a fascinating phenomenon occurs even before birth: fetuses begin to develop an awareness (perception) of the sounds surrounding them, including the pronunciation of their mother tongue.

Newborns possess a remarkable ability to distinguish between different speakers and even different languages. This talent for auditory discrimination is evident in studies like the one by Kishon-Rabin et al. (2016), which demonstrated that infants as young as 4 to 6 months old can differentiate between sounds like /pa/ and /ba/. This ability highlights a crucial stage in language acquisition: phonemic perception. Phonemes are the smallest units of sound in a language that can change the meaning of a word. For instance, in English, the distinction between /p/ and /b/ creates a difference between "pat" and "bat". This constant exposure allows them to unconsciously absorb the intricacies of pronunciation, grammar, and vocabulary. The process is remarkably spontaneous and efficient, happening at roughly the same pace for all children. By a young age, they have mastered the basics of their L1, allowing them to use it creatively and appropriately in various situations. Research by Ohala (2008) suggests that by the age of three, children have not only developed the ability to perceive these phonemic distinctions in their native language but can also produce them accurately.

In contrast, acquiring L2 pronunciation as an adult presents a unique set of challenges. Unlike children with their inherent language-learning abilities, adults must approach L2 pronunciation consciously. They need to actively analyze the sounds, compare them to their native language (L1), and practice producing them accurately. This conscious effort often leads to a slower learning process and can make achieving native-like pronunciation seem impossible.

This difference is heavily influenced by the age of onset (AO), the age at which someone starts learning a new language is highlighted by Long (2013, pp. 5) "The acquisition of an ideal native-like pronunciation in a second language (L2) or dialect is highly likely, but not guaranteed, for individuals who begin learning between 0 and

6 years old. This probability gradually decreases between the ages of 6 and 12, and after puberty, achieving a completely native accent becomes significantly less likely".

Kashiwagi and Snyder (2014) identify several factors contributing to the lack of pronunciation focus in L2/FL classrooms: the perceived unachievability of native-like pronunciation, limited teaching time, and insufficient teacher training in pronunciation methods. These factors prevent teachers from effectively incorporating pronunciation practice and feedback into their lessons.

This review explores methods for training learners to discriminate between similar sounds. This ability helps to prevent both **segmental errors** (incorrect pronunciation of individual sounds) and **suprasegmental errors** (errors in intonation, stress, rhythm etc.,). These training methods often incorporate technology, known as **Computer-Assisted Pronunciation Teaching** or **Training** (CAPT). In this review, the terms are used interchangeably to emphasize the broad applicability of these methods, regardless of the specific learning environment. While both "teaching" and "training" appear throughout the review, a subtle distinction can be made depending on the context:

Teaching: typically refers to a more formal, instructor-led approach, often found in classroom settings.

Training: can encompass both instructor-led and self-paced learning environments, including laboratories or independent practice using technology like CAPT.

Due to my background in Greek language instruction, I'll focus on how these methods can be applied to teach Greek pronunciation in a foreign language setting. Notably, there is a scarcity of research specifically on teaching Greek pronunciation as a foreign language. Therefore, this review will highlight methods that can be effectively implemented in both classroom and self-study environments.

This review begins with a concise overview of the literature on:

- **The development of perception in L1 acquisition**: This section will explore how acquire perceptual and how it affects our ability to learn an L2.
- **The role of L1 interference**: Beginning from the behaviorist Contrastive Analysis Hypothesis (CAH) to Chomsky's Universal Grammar (UG).
- **Models of perception of L2 sounds**: This section will delve into different models that explain how learners perceive sounds in a second language.

Following this literature review, the focus will shift to the research methodology employed in this study. Specifically, we'll explore the search criteria used to identify studies that implement Computer-Assisted Pronunciation Teaching (CAPT) methods.

In conclusion, this review aims to elucidate the efficacy of these methods in enhancing the learners' auditory discrimination skills. This, in turn, will lead to improved pronunciation accuracy (intelligibility).

2. Literature review

The ability to perceive and distinguish sounds (phonemes, intonational cues, foci, etc.,) forms the foundation for human communication and language acquisition. This review delves into the ongoing debate surrounding how infants/ adult L2 develop this remarkable skill in their L1 and L2. I explore three prominent theoretical perspectives:

The Behaviorist View: This perspective emphasizes the role of experience and learning in shaping auditory perception. It suggests that infants begin with a relatively unorganized auditory system and gradually refine their ability to distinguish sounds through exposure to their environment.

The Moderate View: This perspective acknowledges some innate predispositions for sound discrimination in newborns. However, it suggests that these abilities require further refinement through exposure to the specific sounds of their native language (as in Attunement Theory).

The Nativist View (Universal Grammar Theory): This cognitivist theory by Chomsky proposes that infants are born with a pre-wired language faculty, allowing them to perceive all possible sounds in human languages. They then lose the ability to discriminate sounds that are not present in their native language through a process of pruning.

These contrasting viewpoints highlight the complexity of auditory perception development. By examining these theories and the supporting evidence, we can gain a deeper understanding of how infants and adults acquire the ability to perceive and distinguish the sounds.

2.1- The development of perception in L1 acquisition

The ability to perceive and distinguish sounds forms the foundation for human communication and language acquisition. Three prominent theories -

Perceptual Learning Theory, Attunement Theory, and Universal Grammar Theory - offer contrasting perspectives on how this remarkable skill develops in infants. This review delves into these theories, exploring their core tenets and the supporting evidence.

There are three different illustrations on how perception is developed, summarized by Werker (2024) in Fig.1.

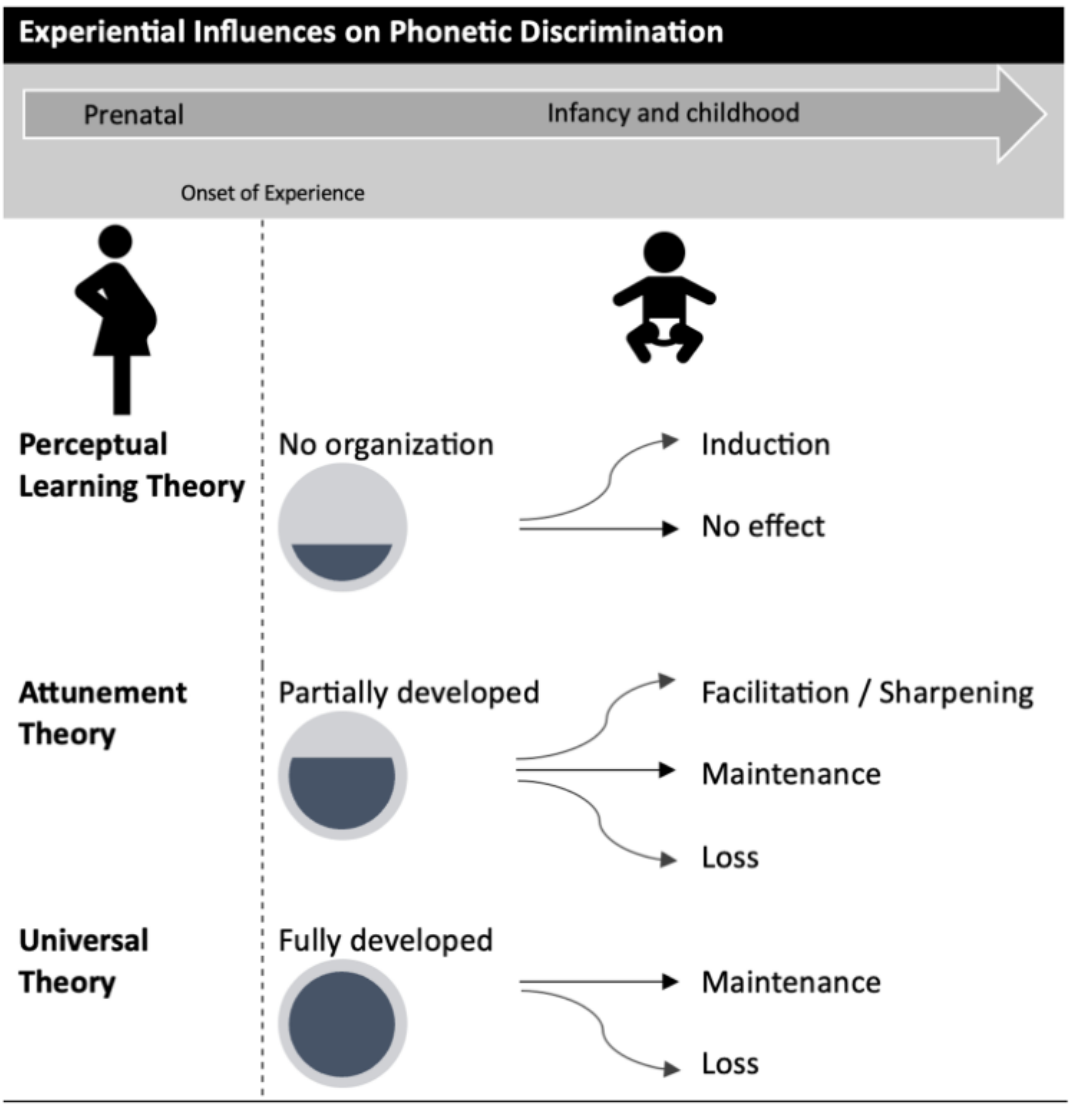

Fig.1. Diagram illustrating the development of auditory perception by Werker (2024)

These three theories paint contrasting pictures of how early auditory perception develops. Perceptual Learning Theory emphasizes the role of experience, while Attunement Theory acknowledges some innate abilities that require refinement. Universal Grammar Theory takes a more nativist approach, proposing a pre-existing language faculty that guides sound perception.

Studies by Werker and colleagues (1980) provided early support for Perceptual Learning Theory. Newborns were shown to be able to discriminate between phonemes present in both their native language and an unfamiliar language, suggesting an initial lack of language-specific perception. However, subsequent research has challenged this view. Neonatal studies by Polka (1991, 1994) demonstrated a preference for native language speech patterns, suggesting some degree of innate bias towards familiar sounds.

Perceptual Learning Theory: A Blank Slate at Birth

Proponents of Perceptual Learning Theory, pioneered by Aslin and Pison (1980), argue that newborns are born with a relatively unorganized auditory system. They posit that infants lack the innate ability to distinguish between different sounds and rely solely on exposure and experience to refine their auditory perception. This theory suggests that the intricate ability to differentiate phonemes, the smallest units of sound that change meaning in a language (e.g., "p" and "b"), is entirely learned through experience with the sounds of their surrounding environment.

Attunement Theory: A Head Start with Refinement

Attunement Theory, championed by researchers like Werker (2024), presents a contrasting view. It acknowledges some pre-existing auditory discrimination abilities in newborns. However, these abilities are considered broad and require further refinement through exposure to the specific sounds of their native language. This theory suggests that infants are biologically predisposed to attend to and learn the sounds they are most exposed to during critical developmental periods.

Universal Grammar Theory: Born with Perfect Pitch, Losing the Symphony

Universal Grammar Theory, a cornerstone of generative linguistics, posits a more radical notion. This theory, as proposed by Chomsky, suggests that infants are born with an innate Universal Grammar, a pre-wired language faculty that allows them to distinguish all possible sounds in human languages. However, as they are exposed to their native language, they lose the ability to differentiate sounds that are not present in that language. This theory implies a process of pruning rather than

learning, with infants specializing their auditory perception based on the sounds they hear most frequently**.**

The debate continues, with researchers investigating the interplay between experience-dependent learning and pre-existing neural predispositions. Further research utilizing neuroimaging techniques and longitudinal studies tracking auditory development may provide a clearer picture.

2.2- The role of L1 interference

This section reviews research on the concept of Interference in the acquisition of L2 phonology, as well as the acquisition of L2 in general. The focus here is on characterizing the main findings and presenting key references related to Interference research.

The idea of Interference in L2 emerged in the 1950s. Trubetzkoy (1939, 1958) proposed that L1 acts as a "sieve" filtering our *perception* of L2. Weinreich (1953) utilized the term "interference" in his classic work on language contact, detailing its impact on various linguistic levels, including phonology, morphology, syntax, and vocabulary. Since the 1950s, numerous theories of phonological acquisition have been developed, yielding diverse and sometimes contradictory results.

One of the most significant hypotheses explaining how learners produce non-native sounds and the extent of their accuracy achievement is the Contrastive Analysis Hypothesis (CAH) proposed by Lado (1957). It states: "Individuals tend to transfer the forms, meanings, and distribution of forms and meanings of their mother tongue (L1) and slang, when they want to express themselves in a foreign language (L2). The same happens when they want to understand the foreign language and its culture" (Lado, 1957: 1).

The strong version of the CAH compares L1 with L2, identifying "similar" and "different" sounds. It posits that elements of L2 like L1 are easier to learn, while those that differ will be more challenging. This allows for predicting potential mistakes L2 learners might make due to negative interference. Fries (1945) emphasizes the importance of lessons based on a scientific comparison of the target language (L2) with the learner's L1. The CAH identifies two types of interference:

- Positive Interference: When an L1 structure resembles an L2 structure, the L2 learner might produce a correct L2 structure (Littlewood, 1984).
- Negative Interference: When an L1 structure differs from an L2 structure, the L2 learner might not produce a correct L2 structure (Van Els et al., 1984).

Examples on Interference are reported by Broselow (1983) who highlights the role of variation in L2 in second language acquisition (SLA). Also, George (1972) found that one-third of the errors in his data could be attributed to negative interference.

However, as the strong version of the CAH failed to predict all language learning difficulties, a weaker version emerged (Wardhaugh, 1979). This version aims to explain or diagnose learner errors that might result from interference (Ellis, 1996). While the CAH can potentially identify possible student mistakes, it cannot predict which phenomena will be more difficult to learn than others.

A less behaviorist illustration has emerged in the 1970s by Eckman's Markedness Hypothesis. Eckman suggests that some linguistic phenomena are inherently more difficult to acquire than others. This principle is particularly relevant in L2 acquisition, especially for phonology. L2 learners must overcome the influence of their native language (L1) and learn the often unfamiliar sound features of L2.

The hypothesis posits that the difficulty of acquiring a new sound in L2 is influenced by the concept of **markedness**. Markedness refers to the relative complexity of a linguistic feature within a language system. Eckman (1987, pp. 60) defines a more marked phenomenon as one whose presence implies the presence of a less marked one, but not vice versa.

For example, in English, voiced consonants like /b/, /d/ at the end of words (e.g., "rob", "bad") are more marked than their voiceless counterparts /p/, /t/ (e.g., "rip", "bat"). This is because voiced consonants are less common in this position. According to markedness hypothesis, learners will have more difficulty acquiring and producing these marked sounds accurately.

Eckman's (1977) Markedness Differential Hypothesis (MDH) builds upon Lado's (1957) Contrastive Analysis Hypothesis (CAH) by incorporating the concept of markedness. MDH proposes that L2 sounds will be difficult to acquire only if they are typologically marked. In other words, sounds that are inherently more complex across languages will be more challenging to learn. Conversely, unmarked sounds should not pose a significant learning obstacle. The MDH makes predictions about difficulty at three key points:

1. Points of differentiation: Sounds that differ between L2 and L1 will be more marked in L2.

2. Degree of difficulty: The degree of difficulty in acquiring an L2 sound corresponds to the degree of its markedness. More marked sounds will be harder to learn.

3. Less marked sounds: Sounds that are different from L1 but not inherently marked will be less difficult to acquire.

While the Learnability and Markedness (LM) framework offers valuable insights into L2 phonology acquisition, it has also faced criticism:

1) Limited Explanation: Archibald (1998) argues that (LM), while helpful in describing general patterns of difficulty, may not fully explain the underlying reasons behind these patterns. He suggests the need for a more fundamental explanation for why certain sounds are more challenging to acquire.
2) Oversimplification: Gass and Selinker (2001) criticize LM for its oversimplification. They argue that simply stating L2 learners follow the rules of their native language (L1) is insufficient. While L1 influence is undeniable, a more nuanced understanding is required to explain why universal principles of markedness apply across languages.

Universal principals as stated in the critique of Gass and Selinker (2001) can be presented through the framework of universal grammar (UG). Pioneered by Noam Chomsky has revolutionized our understanding of language acquisition. It challenged the prevailing belief that children learn language solely through exposure and imitation.

Despite variations across languages, the core principles of UG ensure that all humans possess the basic tools for language acquisition. This theory offers a compelling explanation for the remarkable speed and efficiency with which children learn language.

The applicability of Universal Grammar (UG) to adult second language acquisition (SLA) remains a topic of debate. While some researchers (e.g. Bley-Vroman et al., 1988) suggest that adults retain access to UG principles and parameters, facilitating high-level L2 proficiency, others propose alternative models.

Supporters of the Full Access Hypothesis argue that adults, like children, possess full access to UG during SLA. This allows them to learn a second language by subconsciously setting parameters to match the target language's grammar.

However, some linguists propose a more limited role for UG in adult SLA:

- Indirect Access Model: This model suggests that adults rely on their explicit knowledge of L1 grammar to learn L2 grammar, leading to potential interference from L1 patterns.

- No Access Model: This model posits that adults lose access to UG principles after a critical period in childhood, hindering their ability to fully acquire an L2 grammar.

Cook (1985) further categorizes potential access to UG during SLA through the diagram in Fig.2.

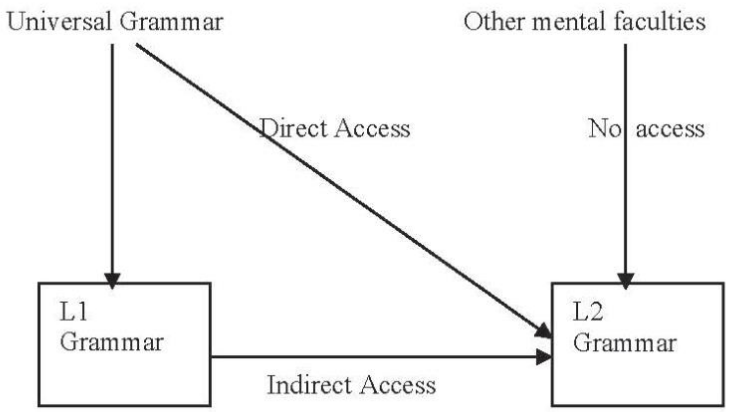

Fig.2. Diagram of UG access by Cook (1985)

Ellis (1994) proposes a processing model that acknowledges both implicit and explicit learning mechanisms in SLA. Adults may initially rely on explicit learning strategies like the Indirect Access Model, but with sufficient exposure and practice, implicit UG-based learning can take over, leading to more natural and fluent L2 production.

2.3- Models of perception of L2 sounds

Adult learners of a new language often exhibit a foreign accent, meaning their pronunciation deviates from that of native speakers. This can be attributed to the influence of their native language (L1) on the production of sounds in the new language (L2) as stated above.

Beyond production challenges, adult learners also face hurdles in *perceiving* the sounds of a new language. Certain sounds in L2 may be absent in L1, making them difficult to distinguish or recognize correctly. For instance, research by Miyawaki et al. (1975) and Yamada (1995) suggests that Japanese speakers (L1 Japanese) struggle to differentiate between the English sounds /l/ and /r/. Experiments involving learners with varying levels of English proficiency confirmed this difficulty. The performance of most Japanese listeners fell significantly below that of native English speakers.

This phenomenon highlights a potential connection between **production** and **perception**. The challenges adult learners face in producing specific sounds in L2 might hinder their ability to perceive those same sounds accurately when spoken by native speakers.

Building on the previous example, Gottfried (1984) found that English learners, regardless of their French proficiency, generally struggled more with distinguishing French vowels compared to German vowels. Similarly, Flege (1995) demonstrated that even experienced Spanish speakers learning English faced challenges differentiating specific vowel pairs like [ɛ/æ] and [a/ʌ], despite having no trouble with other vowel contrasts like [i:/ɪ] and [ʌ/ʊ]. These findings suggest that the difficulty of perceiving L2 sounds is not solely dependent on the learner's overall experience with the language, but also on the specific characteristics of the sounds themselves. Certain sound contrasts may be inherently more challenging to distinguish due to their acoustic similarity or the absence of similar contrasts in the learner's native language.

To gain a deeper understanding of the challenges adult learners face in perceiving L2 sounds, and to explore how these challenges might impact their pronunciation (L2 production), we must delve into the existing theoretical models of L2 speech perception. Examining these models will provide valuable insights and guide our discussion on the intricate relationship between perception and production in second language acquisition.

There are three major models that are highly cited in virtually all articles regarding L2 speech perception. They are Speech Learning Model (SLM), Perceptual Assimilation Model (PAM) and the Native Language Model (NLM). These models are better illustrated in (Derwing and Munro, 2015).

The SLM emphasizes the link between accurate perception and accurate production of L2 sounds. Flege's model builds upon these principles to explain how learners can achieve native-like pronunciation of vowels and consonants in a second language.

Flege's model proposes that adult learners categorize L2 sounds based on their similarity to existing categories in their native language (L1). This categorization process influences both perception and production of L2 sounds.

Similar vs. New Sounds: Learners initially perform better with L2 sounds that resemble sounds in their L1 "similar" compared to entirely new sounds.

L1 Interference: Early in L2 acquisition, L1 categories can impede learning new sounds. Learners may mistakenly believe they already know "similar" sounds when they don't.

French learners struggle with the English /u/ because they categorize it as "similar" to their native /u/, hindering accurate learning. Conversely, the absence of a similar sound in English for the French /y/ helps learners create a new category and perceive it more accurately as reported by Flege.

In another study, French learners also categorize the English /p, t, k / as "similar" to their own /t/, but the English /p, t, k / has a longer closure time. This similarity leads French learners to produce the English /t/ for example with a shorter delay, deviating from native-like pronunciation as reported by Flege & Hillenbrand (1984).

As per Best's (1995) PAM model, which is originally developed to understand how people perceive sounds outside their native language (L1), it offers valuable insights into the challenges faced by second language (L2) learners. PAM proposes that when encountering unfamiliar sounds in a new language, listeners attempt to categorize them based on their similarity to existing categories in their L1 phonological system. This process of *assimilation* influences how learners perceive L2 sounds.

These assimilation strategies can contribute to the foreign accent often observed in adult L2 learners. Understanding the role of assimilation in L2 perception highlights the importance of exposing learners to a wide range of L2 sounds, especially those with minimal L1 counterparts. This can help them develop more accurate categories for new sounds and improve their overall perception of the L2 sound system. These categories of assimilation are distributed into 6 categories as illustrated by Almbark (2012) in Fig.3.

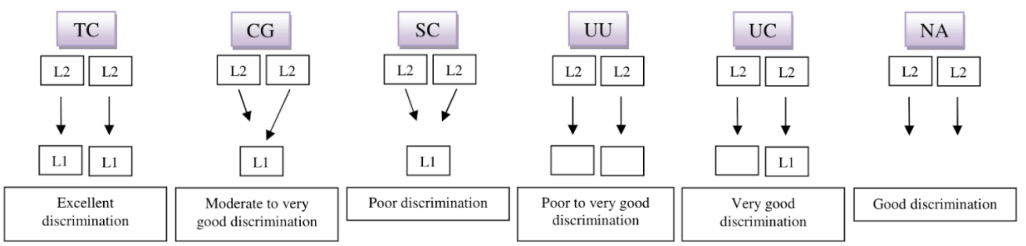

Fig.3. Diagram of assimilation categories as illustrated by Almbark (2012, pp. 48)

While both Flege's Speech Learning Model (SLM) and the Perceptual Assimilation Model (PAM) address challenges in L2 sound perception and production, they differ in their focus and application.

SLM emphasizes the role of individual sounds. It examines how learners categorize both similar and new sounds from the L2 based on their L1 inventory.

While PAM focuses on contrasts between sounds. It explores how unfamiliar sounds in a new language are perceived based on their similarity to existing categories in the learner's native language.

PAM was originally designed to understand perception of sounds outside one's native language, Best & Tyler (2007) recognized its potential application to L2 learning. They developed PAM-L2, which specifically examines whether the way non-native speakers perceive L2 sounds is like how people with no prior exposure perceive unfamiliar sounds. In essence, PAM-L2 investigates if the principles of sound assimilation observed in PAM extend to the process of learning a second language.

Another model that is like SLM and PAM is the Native Language Magnet Model (NLM), proposed by Patricia Kuhl (1994), sheds light on how our native language shapes our perception of sounds throughout life. The theory suggests that our brains categorize sounds into distinct mental spaces based on the patterns present in our native language. These categories act like "magnets," attracting similar sounds and making it easier to remember them.

According to NLM, the sounds we learn in infancy form the foundation for how we perceive all future sounds. These early learned sounds become the "best examples" within each category, influencing how we distinguish similar sounds encountered later. This can make it challenging to differentiate between new sounds, especially if they fall within existing categories in our native language.

The NLM also can offer valuable insights into the challenges of adult second language learning. The model explains why adults may struggle to distinguish between sounds in a new language if those sounds closely resemble existing categories in their native language. In these cases, the brain tends to assimilate the new sound to the closest familiar sound, hindering accurate perception and pronunciation.

3. Methodology

This review aimed to identify relevant research on the impact of technology on L2 auditory perception skills. I conducted a comprehensive electronic search in academic databases as Google Scholar. I then organized the implemented methods into three sub-categories to facilitate analysis and enhance their usability for L2 instructors. This methodological focus ensured we included only studies that investigated the use of computer devices to enhance L2 auditory skills.

4. Search results and application.

4.1- Training through minimal pairs

By minimal pairs I mean the words that differ only by one sound. To train through minimal pairs a method is used, called, High Variability Phonetic Training (HVPT) which is designed to help learners perceive and distinguish sounds in a second language (L2) spoken by different speakers in various environments (Thomson, 2018). In this training learners receive immediate feedback to identify the correct sounds (Logan et al., 1991).

High Variability Phonetic Training (HVPT) is a method designed to help learners perceive and distinguish sounds in a second language (L2) spoken by different speakers in various environments). This method leverages multi-tone recordings of native speakers uttering minimal pairs (word pairs that differ by only one sound) during training (Logan et al., 1991). Learners receive immediate feedback to identify the correct sounds.

Studies have shown HVPT to be effective in enhancing English vowel perception for learners with diverse L1 backgrounds, including Japanese, French, Mandarin, etc., (Lambacher et al., 2005, Iverson et al., 2012; Thomson, 2011, among others).

HVPT can help learners discriminate between L2 consonants that closely resemble sounds in their native language, thereby reducing confusion and improving overall perception (Munro & Derwing, 2006). There have been many studies examining the effectiveness of the HVPT method for teaching chunk perception, such as /l/ and /r/ or /r/ and /w/ to Japanese students (Guion et. al., 2000).

Research is limited, yet some evidence suggests that HVPT can also improve L2 production, leading to better comprehension scores in their study (Bradlow et. al., 1999). Munro & Derwing (2006) recommend this particular method because these errors (among problematic contrasts) have a significant impact on communication

and strongly affected pronunciation and intelligibility ratings. They suggest that pronunciation instruction should prioritize these "high function load" contrasts.

Researchers recommend prioritizing "high functional load" contrasts for pronunciation instruction using HVPT (Munro & Derwing, 2006). These are sound distinctions in L2 that learners commonly struggle with and significantly impact communication.

It's unclear yet if the observed improvements in perception translate to spontaneous speaking situations. Further research is needed to determine the long-term impact on spoken fluency (Sakai & Moorman, 2018). While some studies suggest benefits for production (Bradlow et. al., 1999), more research is required to solidify this aspect of HVPT's effectiveness. Overall, HVPT offers a promising computer-based approach to complement pronunciation instruction, especially in L2 settings with limited exposure to authentic speech (Sakai & Moorman, 2018).

To apply this on Greek, Elsafoury (2023) has used an online phonetic tool that is designed to elicit minimal pairs in 5 languages among them is Greek. Elsafoury (2023) has used the tool to create an auditory training for Egyptian students who learn Greek as a foreign language in Cairo universities. The tool can be accessed for free as illustrated in Fig.4.

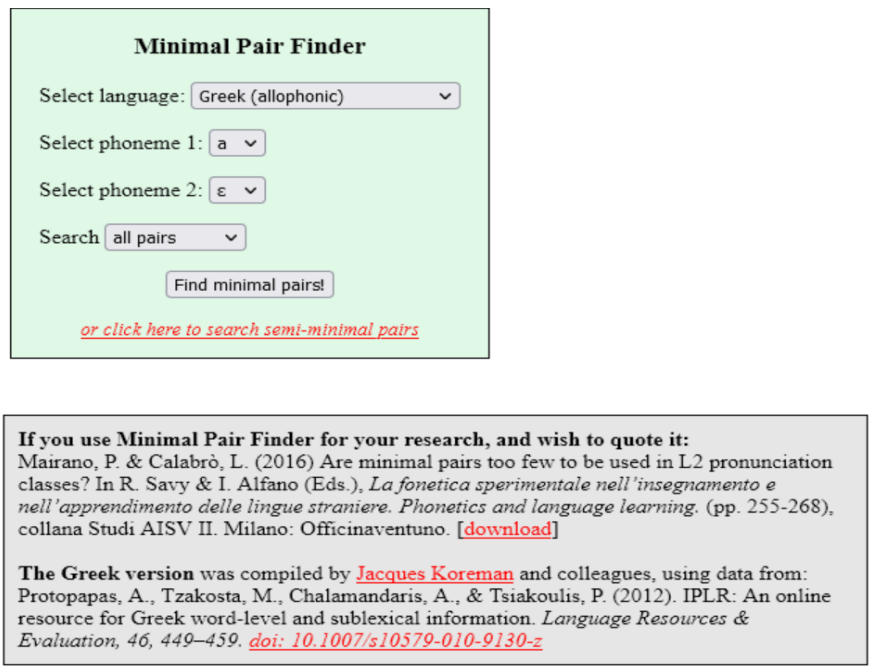

Fig. 4 Minimal pair finder

4.2- Audiovisual reinforcement

In the bibliography (mainly for teaching English pronunciation) there are two types of auditory-visual reinforcement: Interactive and non-interactive. For the purpose of my review I focus mostly on the non-interactive through Praat and SpeakGreek.

Non-interactive audiovisual reinforcements (or aids) can be accoustic related aids or articulatory related aids as reported by (Kröger et. al, 2010). An example for acoustic aids is the use of Praat to show spectrograms of pitch (as in Triassanti 2015), intonation (as in Yoon 2007, Hamlaoui and Bengrait 2016), or even segmental features of vowels and constants (voicing, aspiration, etc.,)

A great example of application for the articulatory aids is a program called SpeakGreek (Nicolaidis et. al, 2015), which is offered by the department of English language at Aristotle University of Thessaloniki. In this web-based program (https://www.enl.auth.gr/speakgreek/) learners can find audiovisual aids for both segmental and suprasegmental features of Greek language. The utterances are produced by female, male and child voices.

In Fig.5 learners are provided with videos illustrating the producing of the sound /i/ in multiple environments through the lips, tongue and air movements.

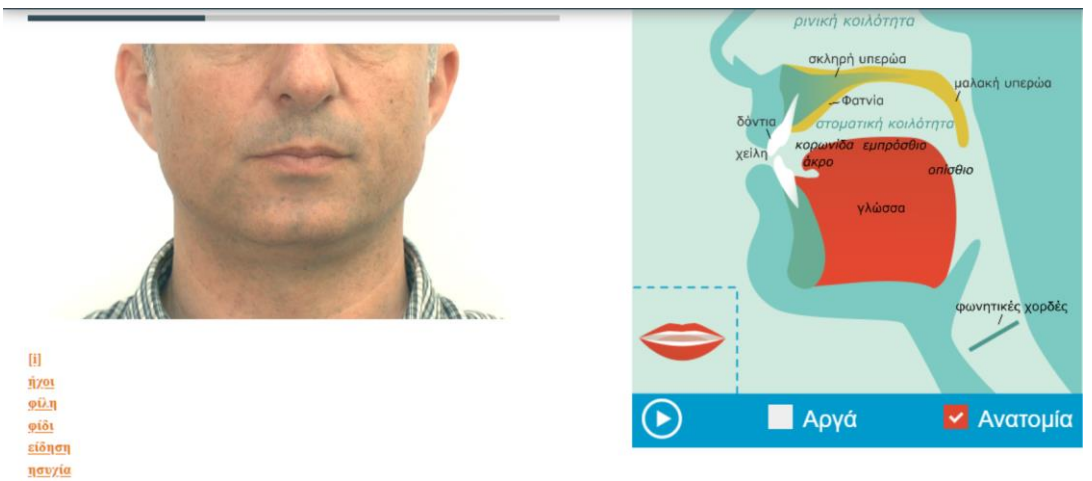

Fig.5. Producing the vowel /i/ by a male speaker as shown on SpeakGreek

By observing these videos, learners gain a deeper understanding of how individual speech sounds (phonemes or segments) like /i/ are produced, contributing to their overall knowledge of segmental elements. The videos also showcase pronunciations of /i/ in different positions (stressed or unstressed, part of onset or

coda) allowing learners to understand the sound's variations. As stated before, learners can choose from a variety of audio recordings featuring female and child voices to observe how the /i/ across different environment. Fig.6 is like Fig.5 where learners can observe the intonation of statements, wh-questions and polar questions. The phrases also produced with multiple foci settings.

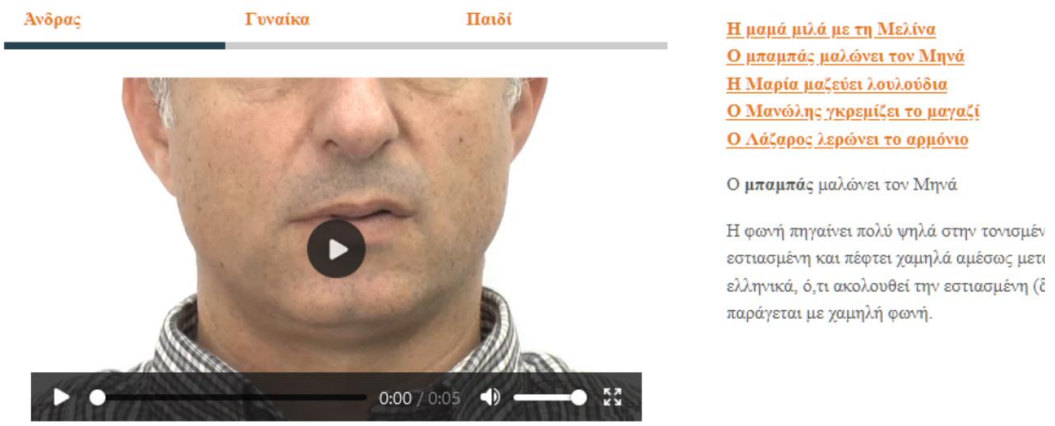

Fig.6. The intonation of a statement produced by a male speaker as shown in SpeakGreek

4.3- Cued Pronunciation Reading

There is a huge interest that underlines the importance of instruction targeting suprasegmental features like rhythm, stress, and intonation (Derwing et al., 1997, among others), there's still a need for more resources that empower learners to practice pronunciation independently.

This review sheds light on the Cued Pronunciation Reading (CPR), a computer-assisted method with promising results for self-directed learning (as reported by Tanner and London 2009). In their study they divided 75 ESL learners into control and treatment groups. The treatment group engaged in 11 weeks of self-directed CPR practice using a computer program. The researchers assessed speech perception and production before (week 1) and after (week 13) the study. The focus was on key suprasegmental features like word stress, pausing and intonation. The results yield significant improvement for the CPR group in perceiving pausing and word stress.

Cued Pronunciation Reading (CPR) has emerged as a compelling method for enhancing learners' perception of pronunciation, particularly suprasegmental features

like rhythm, stress, and intonation. Studies have shown its effectiveness in promoting self-directed learning and improved speech perception. CPR combines visual cues with audio recordings to guide learners in decoding pronunciation. These cues might include symbols or markings overlaid on written text that highlight suprasegmental aspects. By observing these cues while listening to a native speaker pronounce the text, learners actively engage with the suprasegmental features of spoken language. This approach can be particularly beneficial for self-directed learning environments, as learners can practice independently and receive visual reinforcement for proper pronunciation.

5. Limitations and recommendations

In the introduction I stated that there is a significant gap in research and scarcity of studies investigating methods for teaching Greek pronunciation and their impact on learners' perception. This lack of empirical evidence limits the available methods for effective Greek pronunciation instruction.

Moving forward, addressing this challenge of originality in the field of teaching L2 (Second Language) phonology, specifically for Greek, is crucial. Here are some potential avenues for future research: While resources might be limited, studies could analyze the effectiveness of currently employed methods for teaching Greek pronunciation. Researchers could explore and implement innovative approaches specifically designed for Greek pronunciation instruction. This might involve incorporating elements from established L2 phonology methods, such as Minimal Pairs training or auditory discrimination exercises, while tailoring them to the unique sound system of Greek, L1s and high function load contrasts.

Future studies should prioritize learner perception as a key outcome measure (as we argue in the Conclusion). This could involve pre- and post-instruction perception tests to assess improvement in identifying and distinguishing Greek phonemes, pitch cues, etc. By addressing these research avenues, we can move beyond the limited resources currently available and establish a robust foundation of evidence-based practices for teaching Greek pronunciation. This will ultimately lead to improved learning outcomes for students striving to master Greek speech.

6. Conclusion

To conclude this exploration, let's delve into three key findings revealed by these studies. 1) The effectiveness of CAPT methods, 2) Good production precedes good perception, 3) Which pronunciation aspect matter the most.

Studies that examined CAPT methods yielded positive results of teaching where students how to discriminate between single phonemes (as in Wang & Munro, 2004). Other studies have used visual stimuli or a combination of sound and visual stimuli (e.g., Hardison, 2003, Tanner and London 2009). Overall, the use of computers/ mobile devices in pronunciation teaching/ training can be an effective way to help learners improve their pronunciation accuracy. However, it is important to use computer-based methods in conjunction with other teaching methods and to consider the individual needs of learners and environment.

The SLM suggests in its assumptions that there is a correlation between perception and production. According to the SLM, when non-native speakers learn an L2, their speech will eventually reflect the way they perceive sounds in their minds. In other words, if they can accurately perceive the distinctive features of sounds, their production will be affected by this perception. Flege (2003) argued that accurate perception of L2 sounds is necessary for correct production.

Most studies support that improvement will occur at the level of perception rather than production (Munro, 1993, Flege, et al., 1997). Sakai & Moorman (2018), in a review of perception-based studies over the past 25 years, found that there is a correlation between perception and production. They argue that such an instructional approach leads to moderate improvements in perception ($d = 0.92$, $SD = 0.96$). While improvements in production are relatively smaller ($d = 0.54$, $SD = 0.45$). They also argue that the improvements between the two skills (perception and production) are not statistically significant. Their study also found that the improvement in consonant production was greater than that of vowels. However, the study cautions researchers against directly linking perception and production in long-term language development. On the other hand, there are studies that had no effect on perception while contributing to improved production (as in Lacabex and Lecumberri, 2010, Han, 2002).

7. Bibliography

Almbark, R. (2012). *The perception and production of SSBE vowels by Syrian Arabic learners: The foreign language model* (Doctoral dissertation, University of York).

Archibald, J. (1998). *Second Language Phonology*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

articulatory visualization in computer-aided pronunciation training. In A. Esposito, N. Campbell, C.

Best, C. (1995). A direct realist view of cross-language speech perception. In Speech perception and linguistic experience: Issues in cross-language research ed. Strange, W., 171–204. Baltimore: York Press.

Best, C. (1999). Development of Language-Specific Influences on Speech Perception and Production in Pre-verbal Infancy. In proceedings of the International Congress of Phonetic Sciences, 14, 1261–1263. Department of Linguistics, University of California at Berkeley

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (eds.), Language experience in second language speech learning: In honour of James Emil Flege, Amsterdam/ Philidelphia: John Benjamins, 13–34.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Perception & psychophysics*, *61*, 977-985.

Broselow, E. 1984. An investigation of transfer in second language phonology. International Review of Applied Linguistics 22: 253–269.

Cook, V. (1985). Language functions, social factors, and second language teaching, IRAL, 13 (3), pp.177-96.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). John Benjamins Publishing Company.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for "fossilized" learners: Can it help? Applied Language Learning, 8(2), 217-235

Eckman, F. R. (1987). The reduction of word-final consonant clusters in interlanguage, in: A. James - J. Leather (eds.), Sound Patterns in Second Language Acquisition

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. Studies in Second Language Acquisition 18: 91–126.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 229–273). Timonium, MD: York Press.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), Phonetics and phonology in language comprehension and production: Differences and similarities (pp. 319–355). Berlin: Mouton de Gruyter.

Flege, J. E., & Hillenbrand, J. (1984). Limits on pronunciation accuracy in adult foreign language speech production. Journal of the Acoustical Society of America 76: 708–721.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of phonetics*, *25*(4), 437-470.

Fries, C. C. (1945). Teaching and learning English as a foreign language. Ann Arbor: University of Michigan Press.

Gass, S. M. and L. Selinker (2001). Second Language Acquisition. An Introductory Course. (2nd edition) Hillsdale, NJ: Lawrence Erlbaum.

Gottfried, T. L. (1984). Effects of consonantal context on the perception of French vowels. Journal of Phonetics 12: 91–114.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*(5), 2711-2724.

Hamlaoui, N. & Bengrait, N. (2016). Using BetterAccent Tutor and Praat for learning English intonation. *Arab World English Journal*, 3, 99-112.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495-522.

Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, *33*(1), 145-160.

Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, *33*(1), 145-160.

Kashiwagi, A., & Snyder, M. (2014). Intelligibility of Japanese college freshmen as listened to by native and nonnative listeners. *Jacet journal= 大学英語教育学会紀要/大学英語教育学会 編*, (58), 39-56.

Kröger, B. J., Birkholz, P., Hoffmann, R., & Meng, H. (2010). Audiovisual tools for phonetic and

Kuhl, P. K. (1994). Learning and representation in speech and language. Current Opinion in Neurobiology. 4: 812–822.

Lado, R. 1957. Linguistics across cultures. Ann Arbor, MI: The University of Michigan Press. Leather, J. 1999. Second-language speech research: An introduction. Language Learning 49: 1– 56.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, *26*(2), 227-247.

Littlewood, W. (1984). *Foreign and second language learning: Language acquisition research and its implications for the classroom*. Cambridge university press.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874-886.

Long, M. (2013). Maturational constraints on child and adult SLA. In G. Granera & M. Long (Eds), Sensitive periods, language aptitude, and ultimate L2 attainment (pp. 3-41). Amsterdam, The Netherlands: Benjamins

Mairano, P., & Calabro, L. (2017). Are minimal pairs too few to be used in pronunciation classes? *R. Savy, I. Alfano (a cura di), La fonetica sperimentale nell" insegnamento e nell" apprendimento delle lingue straniere, collana Studi AISV, 2*.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. Perception and Psychophysics 18: 331–340.

Munro, M. J. (1993). Production of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. Language and Speech, 36(1), 39e66.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*(4), 520-531.

Nicolaidis, K., Papanikolaou, G., Kainada, E., & Avdelidis, K. (2015). SpeakGreek: An online speech training tool for L2 pedagogy and clinical intervention. In *ICPhS*.

Ohala, D. K. (2008). 1. Phonological acquisition in a first language. In *Studies in Bilingualism* (pp. 19–39). John Benjamins Publishing Company. http://dx.doi.org/10.1075/sibil.36.03oha.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of Acoustical Society of America* 89: 2961–2977.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America 97*: 1286–1296.

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, *39*(1), 187-224.

Segal, O., Houston, D., & Kishon-Rabin, L. (2016). Auditory discrimination of lexical stress patterns in hearing-impaired infants with cochlear implants

compared with normal hearing: Influence of acoustic cues and listening experience to the ambient language. *Ear and Hearing*, *37*(2), 225-234.
synchrony (pp. 337–345). Springer. https://doi.org/10.1007/978-3-642-12397-9_29

Tanner, M., & Landon, M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility.

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Calico Journal*, *28*(3), 744-765.

Thomson, R. I. (2018a). High variability [pronunciation] training (HVPT) A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation*, *4*(2), 208-231.

Thomson, R. I. (2018b). Measurement of accentedness, intelligibility and comprehensibility. In O. Kang, & A. Ginther, (Eds.), Assessment in second language pronunciation. (pp. 11-29) New York: Routledge

Triassanti, R. (2015). Training on word stress of EFL students assisted with Praat program. In A. J. Othaman, N. S. Degang, I. Wiryokusumo, A Susilo, & Mustaji (Eds.), The 6th ICETA International Conference on Education: *"Future education: Shaping intelligent and mannered generation throughout civilization"* (pp. 381–390). University of PGRI Adi Buana Surabaya.

Trubetzkoy, N. (1939). Principles of phonology. Paris: Klincksieck

Trubetzkoy, N. (1958). Grundzüge der Phonologie [Fundamentals of Phonology] (Travaux du cercle linguistique de Prague 7). Göttingen: Vandenhoek & Ruprecht.

Van Els, T., Bongaerts, T., Extra, G., Van Os, C., & Janssen-van Dieten, A. M. (1984). *Applied linguistics and the learning and teaching of foreign languages*. Hodder Arnold.

Vogel, A. Hussain., & A. Nijholt (Eds.*), Development of multimodal interfaces*: Active listening and

Wardhaugh, R. 1970. The contrastive analysis hypothesis. TESOL Quarterly 4: 123–130.

Werker, J. F. (2024). Phonetic perceptual reorganization across the first year of life: Looking back. *Infant Behavior and Development, 75*, 101935.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*, 349-355.

Yoon, K. (2007). Imposing native speakers' prosody on non-native speakers' utterances: the technique of cloning prosody. *Modern English and American Literature, 4(46*), 197-215.