Perception Deception: Exploring the Gap between Self-Perception and Phonemic Perception among Arabic-speaking EFL learners

Faisal Aljasser[1] & Soliman Ismail[1]

[1] Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buraydah, Saudi Arabia

[†]These authors contributed equally to this work and share first authorship.

Correspondence: Faisal Aljasser, Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buraydah, Saudi Arabia. E-mail: jasr@qu.edu.sa

Received: October 24, 2024 Accepted: December 13, 2024 Online Published: March 17, 2025

doi:10.5430/wjel.v15n4p341 URL: https://doi.org/10.5430/wjel.v15n4p341

Abstract

This study explores the relationship between perceived and actual phonemic perception abilities among Arabic-speaking English as a Foreign Language (EFL) learners. It investigates whether learners' perceptions align with their actual performance and whether they tend to underestimate or overestimate their phonemic abilities. Fifty-eight participants, native speakers of Qassimi Arabic, rated the perceived difficulty of English vowels and completed a vowel perception task. Results reveal a significant discrepancy between perceived and actual abilities, with most participants underestimating their phonemic perception skills. A weak positive correlation between perception and performance suggests that learners' self-assessments may not reliably reflect their actual abilities. Findings provide further empirical evidence of the Dunning-Kruger effect (Dunning, 2011) and extend such evidence to include Arab EFL learners' perceived phonemic abilities. Theoretical, epistemological, and pedagogical implications of the study are discussed, including a call for less reliance on learners' self-perceptions in L2 research and instruction in favor of objective performance measures.

Keywords: controlled study, EFL vowels, English language learners, metacognitive awareness, oral language skills, perception, perceived difficulty, quantitative analysis

1. Introduction

The ultimate goal of learning a second language (L2) often involves attaining a high level of linguistic competence, which encompasses both linguistic accuracy (Chomsky, 1966) as well as the ability to use the language effectively in diverse contexts (Hymes, 1966). A fundamental question in the field of second language acquisition (SLA) revolves around why certain learners are more successful than others in reaching this goal (Davies, 2003). Various linguistic, psychological, and social frameworks have sought to answer this question from different perspectives (Saville-Troike, 2016), with some psychological models highlighting the influence of individual differences on the degree of language learning success (Lightbown & Spada, 2021). One psychological approach that focuses on individual differences in cognitive language learning is the information processing model (IPM). This model posits that language learning involves a progression from controlled, conscious language processing to automatic, unconscious processing through repeated practice and exposure (Dekeyser, 2001; McLaughlin & Heredia, 1996). From an information processing standpoint, the cognitive and metacognitive learning strategies that learners employ to process language are significant factors that contribute to varying levels of success in L2 learning (O'Malley & Chamot, 1990).

Cognition involves the essential mental processes that L2 learners employ to notice, perceive, and understand new language experiences, categorize and integrate new knowledge with existing mental representations of language, store this new knowledge in the brain, replace old knowledge structures when needed, utilize internalized linguistic information for language tasks, and thus develop linguistic skills. Successful listening comprehension, for example, depends on how successful individuals are in processing language input throughout the three interconnected cognitive stages of comprehension, i.e., speech perception, parsing, and utilization (O'Malley & Chamot, 1990). Unlike Krashen (1982), who claims that comprehensible input is enough for language acquisition, information processing theorists emphasize that learners must consciously notice differences between L2 input and their existing native language (L1) structures for successful language acquisition (R. W. Schmidt, 1990). Some language learners are more adept at using appropriate learning strategies during this cognitive process than others (R. Schmidt, 2001).

Metacognition, on the other hand, refers to learners' awareness and control over their own cognitive learning processes. Schraw and Moshman (1995) identify three types of metacognitive knowledge: declarative, procedural, and conditional. Declarative knowledge involves learners' perceptions and beliefs about language learning, including their own abilities, strategies, and the tasks they are engaging with. Procedural knowledge refers to learners' awareness of how to apply learning strategies effectively. Conditional knowledge encompasses learners' understanding of 'when and why' to use particular strategies (Schraw & Moshman, 1995). Individual learners not only differ in their metacognitive knowledge repertoires but also in how they utilize such awareness to 'regulate' their own language

learning—that is plan, monitor, and evaluate their learning strategies and outcomes (Brown et al., 1983; Flavell, 1979). Proper metacognitive awareness enables students to rectify their performance when necessary (Huff & Nietfeld, 2009). Research has shown that learners who effectively use metacognitive learning strategies tend to be more successful (e.g., Bercher, 2012). They are better at setting realistic goals, self-assessing their progress, and refining their learning strategies. In addition, understanding metacognitive strategies involved in language learning is crucial for developing effective teaching methodologies (C. C. M. Goh & Vandergrift, 2022). For a taxonomy of cognitive and metacognitive strategies in SLA, see O'Malley and Chamot (1990) and Oxford (1990).

In recent years, there has been a growing research interest in metacognitive awareness in SLA (see Wenden, 1998 for a review). In particular, more researchers than ever have been investigating language learner perceptions—the individual views, interpretations, and attitudes that language learners hold regarding various aspects of their learning experiences, abilities, and strategies enabling them to monitor and evaluate their learning progress. A Google scholar search using the key phrase, 'perceptions of language learners' yielded more than 3.12 million results, with over 20,400 papers published in the first nine months of 2024 alone. SLA studies have explored language learners' perceptions in relation to their own abilities, the difficulty of language tasks, the effectiveness of teaching methods, the relevance of materials, and the feedback they receive. Researchers have observed that learner perceptions are shaped by personal experiences, cultural background (Schulz, 2001), prior knowledge, and the learning environment, significantly influencing motivation, engagement, and overall success in language acquisition (for a review, see Wesely, 2012).

Assuming that we accurately measure learner perception using valid and reliable instruments, understanding learner perception is crucial for effective language learning and teaching (C. C. M. Goh & Vandergrift, 2022) and ultimately helps learners succeed in their learning endeavors. If learners' perceptions do not align with their actual abilities, this potential mismatch can lead to either overconfidence or undue pessimism, both of which can hinder their learning progress (Dunning et al., 2004). Since learner perception, like many nonlinguistic variables, cannot be directly observed, SLA researchers tend to rely on indirect methods to gauge it (e.g., C. Goh, 1997), most notably through questionnaires, self-evaluations, and other self-report measures (Ellis, 1994; Gass et al., 2020).

The reliance on self-report instruments in studies of learner perception within second and foreign language research may be problematic due to the unexamined theoretical assumptions underlying these methods. For instance, most studies assume that language learners can accurately and objectively assess their own language abilities. This presupposes a direct correlation between learners' perceptions of their linguistic abilities and their actual performance. However, this assumption is rarely scrutinized despite evidence suggesting that learner self-reports can be biased and may not accurately capture the dynamic nature of metacognitive processes (Dörnyei & Dewaele, 2023; Gass et al., 2020; Mackey & Gass, 2022). Moreover, learners often tend to either underestimate or overestimate their abilities (e.g., Dunning, 2011).

One significant research area in the context of SLA is the validity of L2 learners' perceptions of the difficulty of linguistic structures in general and phonemic perception in particular. For learners' perception to guide the language learning process effectively, it should be at least demonstrated that perception reflects some aspects of linguistic performance. If there is no correlation between perception and performance, then the role of learner perception in language learning may need to be reconsidered. Unfortunately, existing studies exploring this relationship provided inconsistent conclusions largely due to methodological limitations. These limitations include reliance on indirect measures of linguistic performance, such as course grades, rather than assessment of specific linguistic structures (e.g., Cheng, 2005), potential post-task assessment bias in measuring perceived task difficulty, and potential interaction effects of other study variables (e.g., Ortega et al., 2022; Wu, 2019). This leaves a significant gap in our understanding of whether learners' perceived difficulty truly reflects their actual linguistic performance. This study specifically focuses on the perception of vowel sounds in EFL by Arabic-speaking learners. We hypothesize that if learner perception of linguistic ability truly reflects actual linguistic ability, Arabic-speaking EFL learners' perceived phonemic perceptions will positively correlate with scores on a phonemic perception task.

The current study aims to address this critical gap by adopting a well-controlled study design that directly measures both perception and performance to examine the theoretical assumption that language learners' perceptions of their linguistic abilities are valid measures of these linguistic abilities. To be more specific, the study empirically examines the relationship between perceived difficulty and actual performance in the context of English vowel perception among Arabic-speaking EFL learners. This approach not only has the potential to inform educational practices by emphasizing the importance of aligning learners' self-assessments with their actual linguistic capabilities, but also contributes to a more nuanced understanding of perception in SLA. It also lays the groundwork for more informed epistemological practices in future research on learner perception.

2. Perception of L2 Phonemes

Different models of speech perception have hypothesized how L2 perception takes place in relation to L1 inventory. One of the most influential models of cross-language speech perception is the Perceptual Assimilation Model (PAM) (Best, 1995), which was more recently revised to best account for L2 learners' perception (Best & Tyler, 2007). PAM maintains that the quality of discriminating non-native phonemic contrasts relies heavily on how these non-native contrasts are assimilated to native phonological categories. Similarly, the Speech Learning Model (Flege, 1995) proposes that L1 and L2 systems share the same phonological space. Accordingly, non-native speakers are less likely to notice and thus assimilate to L1 categories the L2 segments that are phonetically similar to these L1 categories. Consequently, these similar L2 categories will be harder to learn. On the other hand, perceptually distinct L2 categories will be easier to learn because they are more likely to be noticed and thus less likely to be mistakenly assimilated to L1 categories as they occupy

non-L1 spots in the phonological space. From an information processing perspective, some language learners are more adept at noticing cross-phonemic differences during this cognitive process than others (R. Schmidt, 2001).

Abundant empirical evidence suggests that L2 phonemes not available in L1 can cause misperceptions in the recognition of L2 words. For example, Broersma and Cutler (2011) investigated in a series of multiple experiments how minimal pairs containing non-native vowel contrasts (e.g., *cattle* vs. *kettle*) are perceived by Dutch-speaking L2 English learners. They found that non-native vowel contrasts were misperceived and resulted in the activation of non-target words. Similarly, Evans and Alshangiti (2018) investigated the perception of British English vowels and consonants in quiet and noise by native Saudi Arabic learners of English from different proficiency levels. Their findings indicated that for all learners the identification of consonants was better than the identification of vowels in both quiet and noise conditions. It was also found that even advanced learners had problems perceiving English vowels that had no L1 Arabic counterparts. In the context of the current study, the investigation of vowel perception is consistent with previous findings which show that vowels are more difficult to perceive than consonants. Therefore, the investigation of vowel perception and its link to perceived difficulty by EFL learners is expected to be more revealing than consonant perception. The next section presents a comparison between the vowel inventory of Arabic and English.

2.1 Vowel System in Arabic and English

Arabic language enjoys a diglossic situation in which Arabic speakers use two language varieties simultaneously (C. A. Ferguson, 1959; Holes, 2004). On the one hand, Modern Standard Arabic (MSA) is used in rather formal situations and in written language. On the other hand, the local variety is used in everyday spoken communications. Therefore, it is essential to examine L1 influence from both varieties when investigating Arabic speakers. Participants in the current study speak Qassimi Arabic (QA) as a local dialect, which is a subset of Najdi Arabic and is spoken in the Qassim region in the center of Saudi Arabia. Therefore, the vowel systems of both MSA and QA are considered.

Arabic vocalic system is relatively limited. MSA has three short vowels /i, a, u/ and their long counterparts /i:, a:, u:/. The two diphthongs /ay/ and /aw/ complement MSA's vocalic system (Ryding, 2005; Watson, 2007). QA—as a subdialect of Najdi Arabic—vocalic system uses the same short and long vowels as MSA. However, in QA, the diphthongs /ai/ and /au/ have been monophthongized to /e:/ and /o:/, respectively (Ingham, 1994). In contrast, English has a complex vocalic system. Received Pronunciation (RP), for example, has a total of 20 vowel sounds. These include seven short vowels /ɪ, e, æ, ʌ, ɒ, ʊ, ə/, five long vowels /i:, ɑ:, ɔ:, u:, ɜ:/, and eight diphthongs /eɪ, aɪ, ɔɪ, əʊ, aʊ, eə, ɪə, ʊə/ (Roach, 2009).

3. Literature Review

A few studies have attempted to establish a link between learners' perceptions and actual linguistic performance. For example, Cheng (2005) investigated the associations between anxiety, self-perceived competence and actual course performance in EFL listening comprehension. Cheng asked 23 EFL learners to rate their listening comprehension anxiety and proficiency level on a five-point scale ranging from five (very high) to one (very low). Course grades were also collected. Cheng found out that while higher anxiety was associated with lower course grades, self-perceived competence had a stronger negative association with anxiety than grades. These results demonstrate that self-perceptions might be guided by some psychological factors such as anxiety and may not necessarily reflect actual performance. However, although Cheng's study shows that self-perception may not guide linguistic performance, its methodological design makes it difficult to draw generalizable conclusions. That is mainly because linguistic performance here is based on course grades rather than an assessment of a specific linguistic structure.

Another area where the effect of perceived difficulty was investigated is translation. Wu (2019) examined the link between text characteristics, perceived difficulty, and task performance in sight translation. Participants were asked to sight-translate six texts which differed in their syntactic, lexical, and discoursal properties. Task performance was measured in terms of both accuracy and fluency. After translating each text, the participants were asked to rate their perception of task difficulty using five questions on a six-point Likert scale. Wu (2019) found out that perceived difficulty positively correlated with task performance. However, a closer look at the data collection procedure may allow for other interpretations. Given that participants rated perceived difficulty after the end of the task and not before, their ratings may not have reflected their perception of difficulty but rather their awareness of the mistakes they made during the task.

In another study, Ortega et al. (2022) examined whether non-native learners' perceptions of the comprehensibility and accentedness of their own L2 speech match native and non-native listeners' evaluations. They also aimed to determine if L2 learners' self-assessment of comprehensibility and accentedness was related to their perception of task performance. A story-telling task was used to elicit the oral narrative. Participants were asked to complete a task performance questionnaire immediately after completing the speaking task. The results showed that non-native learners' perceptions and native/non-native listeners' evaluations were only moderately related for comprehensibility but not for accentedness. Moreover, there was no correlation between L2 learners' self-assessment and their perception of task performance. However, the general trend was that L2 learners who had a strong accent and were the least comprehensible to native/non-native listeners tended to over-rate themselves, and vice versa. These results (see also Saito et al., 2020) are consistent with the Dunning-Kruger effect (Dunning, 2011), which "describes the tendency for unskilled performers to overestimate their ability relative to external assessments. Skilled performers, who are often equally inaccurate, instead tend to underestimate their performance" (Saito et al., 2020, p. 1).

So far, we have argued that most perception studies in L2 learning have not attempted to find out if there is an association between

learners' perceptions and linguistic performance. These studies were more reliant on a theoretical presupposition of the existence of a perception-performance link. On the other hand, those few studies which examined such link provided inconsistent evidence mainly because of methodological inadequacies. For example, in Cheng's (2005) study performance was measured through final course grades. Unlike standardized language proficiency tests, course grades may not precisely measure listening comprehension proficiency. Moreover, listening comprehension is a major skill that involves various subskills (Field, 2009). Therefore, it is unlikely that L2 learners fully understood the complexity of listening comprehension when they were estimating their perception of their proficiency. Similarly, the other reviewed studies suffer from methodological shortcomings. In both studies by Wu (2019) and Ortega et al. (2022) perceived difficulty was checked after performing the task. Therefore, this may reflect participants' actual realization and awareness of the mistakes they made in the task rather than their perceptions. Moreover, both studies targeted complex linguistic tasks and concepts (i.e., sight translation, comprehensibility and accentedness) that cannot be perceived systematically by learners. For example, whereas some learners may perceive comprehensibility as more dependent on clear pronunciation, others may relate it more to correct and rich vocabulary usage (Trofimovich et al., 2016). Other researchers have pointed out similar critiques. Wesley (2012, p. 103), for example, argues that:

> Associating or correlating two learner attributes, such as learner self-efficacy and high levels of proficiency, does not provide proof that self-efficacy produces high proficiency. The possibility cannot be ruled out that students who have higher proficiency, perhaps as a result of an unknown or unmeasured variable like access to better learning strategies or members of the target community, then have higher self-efficacy.

Consequently, the aim of the current research is to investigate the potential association between learners' perceptions and their linguistic performance through a meticulously designed research approach that can more effectively examine this connection.

4. The Present Study

Considering the limitations in earlier studies, the current study attempts to provide empirical evidence for the availability of the perception-performance link, or lack of it, using a linguistic structure that can potentially tap into this link directly. Here we use the vowel system of EFL. Specifically, the study endeavors to answer the following research questions:

1. Do Arabic-speaking EFL learners accurately perceive their phonemic perception abilities?
2. If not, do Arabic-speaking EFL learners tend to underestimate or overestimate their phonemic perception abilities compared to their actual performance?
3. Is there a correlation between individual Arab-speaking EFL learners' perceptions of their phonemic ability and their actual phonemic perception?

Methodology

4.1 Participants

This study engaged a group of fifty-eight male undergraduate students, all native speakers of Qassimi Arabic, a dialect of Najdi Arabic prevalent in the Qassim region of central Saudi Arabia. These participants, voluntarily involved in the study, were enrolled in the English Language Program at a university in Saudi Arabia, specifically at the second level of an eight-level program spanning four years. The age range of these learners was between 18 and 20 years.

4.2 Tasks

Two tasks were used. The first was a perceived difficulty task in which participants rated their perceived difficulty of phonemically perceiving EFL vowels. A perceived difficulty test (PDT) was used to examine the perceived difficulty of aurally presented RP vowels. The PDT required participants to rate the difficulty of phonemic perception of a given vowel on a five-point scale, with 1 indicating 'very easy' and 5 indicating 'very difficult.' Scores were later reversed to accommodate the only possible logical *a priori* assumption that students who perceive the task more difficult are expected to score lower on the objective test and vice versa.

The second test was a vowel perception test (VPT). In this test, participants listened to the same vowel sounds as in the PDT but in the context of /h/-V-/d/ words (e.g., had, heed). Their task was to identify each aurally presented word from a written set of alternatives, with one being the correct answer (e.g., had, heed, hard, heard). To prevent unfamiliarity with these words among our EFL participants from influencing the results, a common rhyme word was included in each choice, following the approach of Evans and Alshangiti (2018) (e.g., hid as in kid).

4.3 Stimuli

PDT. Eleven RP vowels served as stimuli for the PDT. These included the six short vowels /ɪ, e, æ, ʌ, ɒ, ʊ/ and the five long vowels /iː, ɑː, ɔː, uː, ɜː/. A female native speaker of Standard Southern British English (SSBE) recorded these vowels. The vowels were recorded using a high-quality microphone on to digital-audio-tape at a sampling rate of 44.1 kHz and were saved to a computer disk.

VPT. The same 11 vowels from the PDT were used in the VPT within the context of CVC (i.e., /h/-V-/d/) words (e.g., had, heed). Eleven /h/-V-/d/ words comprised the VPT. These words were embedded in the carrying sentence 'I say a ____.' The sentences were recorded by the same female native speaker of SSBE using a high-quality microphone on to digital-audio-tape at a sampling rate of 44.1 kHz and were saved to a computer disk.

4.4 Procedure

The experiment began with the PDT followed by the VPT. This order was chosen to ensure that the PDT accurately reflected participants' long-term perceived difficulty of the vowels, rather than being influenced by their performance in the VPT. Both tests took place in a language lab with a capacity of 30 seats, leading to the participants being divided into two groups. The aural stimuli were presented through the lab speakers, and data collection was managed through Blackboard LMS.

In the PDT, participants were instructed to listen to a number of English vowels in isolation and to rate the difficulty of phonemic perception (i.e., identification) of each vowel on a five-point scale, with 1 indicating 'very easy' and 5 indicating 'very difficult.' Each vowel was presented three times, and the scale was displayed on the screen during the presentation of the vowel.

Following the completion of the PDT, the VPT was promptly commenced. Participants were instructed to listen to a number of English words embedded in the carrying sentence, 'I say a ____.' Their task was to identify each aurally presented word from a written set of alternatives, with one being the correct answer (e.g., had, heed, hard, heard). Each choice was listed with a common rhyme word (e.g., *hid* as in *kid*). Each word was presented three times.

5. Results

The present study explores the relationship between perceived and actual phonemic perception among Arabic speaking EFL learners. The study further examines if these EFL learners under perceive or over perceive the difficulty of phonemic perception of EFL vowels. A bootstrapped Spearman's rank correlation test and an independent samples *t*-test were utilized to analyze the study data.

As seen in Table 1, which provides descriptive statistics for both variables in the study, the mean score for the actual phonemic perception of the participants was 3.73 (SD = 0.985). The bias-corrected and accelerated (BCa) bootstrapped 95% confidence interval for the mean ranged from 3.48 to 3.973. This suggests a moderate level of phonemic recognition ability among participants. In contrast, the mean score for perceived phonemic perception was significantly lower at 2.315 (SD = 0.555); the 95% confidence interval for the mean ranged from 2.177 to 2.46. This suggests that participants generally underestimated their phonemic perception abilities.

The standard deviations for the two variables (i.e., .985 and .555 respectively) indicate substantial variability in both actual and perceived scores, with actual scores showing greater variability than perceived scores. This suggests that participants' abilities and their perceptions of those abilities vary widely, which might be due to differing levels of self-awareness or objective proficiency among learners.

Table 1. Means and Standard Deviations with Bootstrapped Confidence Intervals of Participants' Actual Phonemic Perception and Perceived Phonemic Perception

				Bootstrap[a]			
						BCa 95% Confidence Interval	
			Statistic	Bias	Std. Error	Lower	Upper
Pair 1	Actual Phonemic Perception	Mean	3.730	.000	.127	3.480	3.973
		n	58				
		Std. Deviation	.985	-.012	.0838	.822	1.114
		Std. Error Mean	.129				
	Perceived Perception	Mean	2.315	.001	.0716	2.177	2.460
		n	58				
		Std. Deviation	.555	-.006	.0438	.477	.621
		Std. Error Mean	.0728				

a. Bootstrap results are based on 10,000 bootstrap samples

A BCa bootstrapped Spearman's rank correlation coefficient analysis was computed to assess the relationship between the participants' perceived phonemic perception and their actual phonemic perception. To be more specific, the statistical analysis sought to investigate whether perceived difficulty of EFL vowels correlated with actual phonemic perception of these vowels by Arabic speaking EFL learners or not. Spearman's rank correlation was selected for two main reasons: (a) Spearman's ρ does not require data to be normally distributed, but in the meantime (b) it is adept at identifying monotonic relationships, regardless of whether these relationships are linear or non-linear. Since visual examination of the study data suggested that the data showed some deviation from normality and the relationship between the two variables was not linear, Spearman's rank correlation was deemed the most appropriate statistical tool for capturing the nuances of the relationship between perceived and actual phonemic perception in this study, ensuring that the analysis remains valid even under the constraints of non-normality and non-linearity.

The results, as presented in Table 2, reveal no strong evidence of a consistent monotonic relationship between the two variables of the study. Specifically, Spearman's rank correlation indicates a non-significant, weak positive correlation between learners' perceived phonemic distinction abilities and their actual performance on phonemic perception tasks, $\rho(56) = .17$, $p = .197$, 95% CI [-.098 - .422]. Only a negligible 2.96% of the variability in actual phonemic perception scores can be explained by the variability of perceived ability. This suggests a pronounced disconnect between the students' perceived and actual phonemic perception. The broad BCa confidence interval (-.098 to .422) confirms that the suggestion that the true correlation could be negligible or non-existent. This highlights the uncertainty in the relationship between perceived and actual phonemic perception.

Table 2. Means, Standard Deviations, and Spearman's Rank Correlation Coefficient with Bootstrapped Confidence Intervals Between Actual Phonemic Perception and Perceived Phonemic Perception

				Perceived Phonemic Perception				
						Bootstrap[a]		
Variable	*n*	*M*	*SD*	Spearman's ρ(56)	*p* (2-tailed)	Bias	Std. Error	BCa 95% Confidence Interval
Perceived Phonemic Perception	58	2.315	.555					
Actual Phonemic Perception	58	3.730	.985	.172	.197	-.003	.138	[-.098 - .422]

[a]. Bootstrap results are based on 10,000 bootstrap samples.

While the correlation analysis we conducted indicated that there is no statistically significant correlation between EFL learners' metacognitive awareness of their phonemic perception and their phonemic perception performance, it does not answer the question of whether EFL learners tend to under perceive or over perceive the difficulty of phonemic perception of EFL vowels. To pursue this inquiry, a BCa bootstrapped paired samples *t*-test was conducted to compare the actual phonemic perception scores and perceived phonemic perception scores.

The *t*-test results (see Table 3) reveal a statistically significant difference between the participants' perceived and actual phonemic perception scores, $t(57) = 10.080$, $p < .001$. The mean difference was 1.415 (95% CI [1.134 - 1.696], SD = 1.070), indicating learners consistently underestimated their phonemic perception abilities. The effect size, calculated using Cohen's *d*, was 1.324 (95% CI [.967 - 1.676]). This large effect size suggests a substantial practical significance in the discrepancy between perceived and actual phonemic abilities and points to a pervasive underestimation that might influence learners' language learning strategies and confidence.

Table 3. Mean Difference between Actual Phonemic Perception and Perceived Phonemic Perception

	Paired Differences						Effect Size		
	M	*SD*	*SEM*	95% Confidence Interval of the Difference	*t*(57)	Sig. (2-tailed)	Standardizer[a]	Point Estimate	95% Confidence Interval of the Effect Size
Actual Phonemic Perception - Perceived Phonemic Perception	1.415	1.070	.140	[1.134, 1.696]	10.080	.001***	1.324	1.324	[.967, 1.676]

[a] The denominator used in estimating the effect sizes is the sample standard deviation of the mean difference adjusted by the correlation between measures.

*** $p < .001$.

The scatterplot in Figure 1, which visually illustrates the relationship between actual and perceived phonemic perception, confirms the significant gap between self-assessed and actual abilities. It further helps understand the patterns and nuances in the statistical analyses reported above. In addition to portraying a negligible positive correlation, the broad dispersion of data points around the line of best fit in the scatterplot and the large standard deviation, namely 1.070, of the 1.415-point mean difference between actual and perceived phonemic perception indicate significant variability in how the learners perceived their phonemic abilities relative to their actual performance. Some clusters of data points indicate two distinct subgroups within the sample that might possess varying levels of metacognitive awareness of their perception skills. The top part of the scatterplot shows that, except for a single student who was able to accurately estimate his actual phonemic perception, participants scoring 2.73 points or higher (84.48% of the sample) on the actual perception task notably not only underestimated their performance (by an average of 1.77 points) but also showed a huge variability doing so (i.e., from 1.36 to 3.45). Even the participants who had a perfect score of five points underestimated their capabilities by at least 1.55 points. In contrast, participants scoring below 2.73 points (13.79% of the sample) tended to overestimate their abilities, albeit with less variability (mean overestimation: 0.56 points). This shift from underestimation among proficient learners to overestimation among less proficient ones is not captured by the t-test and warrants further investigation.

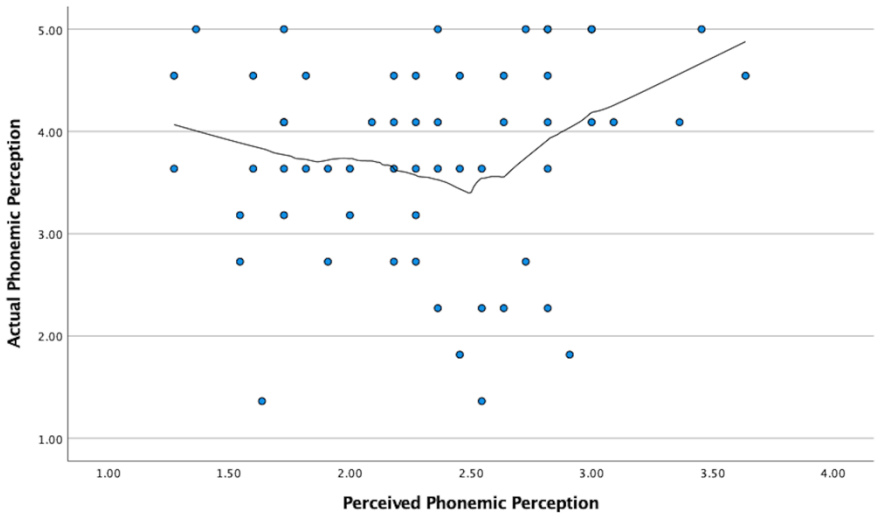

Figure 1. Scatterplot of the Correlation between Actual Phonemic Perception and Perceived Phonemic Perception

6. Discussion

This study investigated the relationship between perceived and actual phonemic perception abilities among EFL learners to assess if learners' perceptions of linguistic ability could substitute for objective assessment of linguistic ability. Two research questions examined whether learners accurately perceived their phonemic abilities and whether they tended to underestimate or overestimate them. Results from a paired samples *t*-test revealed a statistically significant discrepancy between perceived and actual abilities, with participants, as a group, underestimating their phonemic perception by an average of 1.415 points with a large effect size (d = 1.324), indicating limited awareness of their competencies.

The third research question assessed whether a significant correlation existed between perceived and actual abilities. Correlation analysis indicated no statistically significant relationship; learners who rated themselves higher were not necessarily better at phonemic perception, nor were those who rated themselves lower necessarily worse. Overall, the weak correlation does not support the hypothesis that EFL learners' perceptions of their phonemic perception ability is a reliable predictor of actual phonemic perception ability. This reaffirms the finding from the *t*-test and calls into question the over-reliance on learners' perceptions of linguistic ability in lieu of objective assessment.

Further analysis of the scatterplot revealed distinct trends: Higher-performing learners tended to underestimate their abilities, while lower-performing learners overestimated theirs. These findings differ from Stan's (2024) study, which reported a positive correlation between self-assessments and language proficiency scores. The discrepancy may stem from Stan correlating self-assessments with broad proficiency levels (e.g., A1, B2) rather than raw scores, potentially inflating the correlation. Additionally, rather than assessing specific linguistic skills, Stan used the CEFR self-assessment scale, which employs generalized 'Can-Do descriptors.' While generalized tools such as the CEFR descriptors are useful for broader skills, assessing more specialized linguistic abilities requires more targeted and clear measures in both self-assessment and external evaluations. The direct comparison of raw self-assessment and criterion-based test scores in our study provides a more nuanced understanding of learners' capabilities.

In contrast, our study employed a more focused self-assessment tool that targeted phonemic perception, a specific linguistic skill. We believe this approach allowed for a more precise and valid comparison between learners' self-assessments and actual performance. Supporting our view, Ross found that correlations between students' self-assessments and criterion tests were often low; Janssen-van Dieten (1989; 1998) reported that direct-question assessments were better predictors of criterion scores than broader questionnaires. By focusing on phonemic perception, a skill that is less intuitive and requires greater metacognitive awareness, our study illustrates the challenges learners face in self-assessing areas demanding precision. Furthermore, although Stan (2024) reported moderate correlation effect sizes between self-assessments and performance (R^2 = 46.24% for Listening, 25% for Reading), her results reveal that a significant portion of the variance in learners' performance—53.76% for Listening and 75% for Reading— was unexplained, suggesting that factors other than self-assessment accuracy played a major role in determining performance. We argue that the specificity of our self-assessment tool contributes to a more accurate reflection of learners' ability to assess themselves.

Interestingly, although Stan (2024) found a positive correlation between her participants' self-assessments and their performance, she noted that participants generally underestimated their writing skills and overestimated their reading abilities. This further supports the idea that students do not always accurately assess their own language skills.

Our findings also diverge from those of Cheng (2005), Ortega et al. (2022), and Wu (2019). Cheng's reliance on final course grades to measure linguistic performance may not accurately represent students' true abilities, highlighting a limitation in methodology. The discrepancies with Ortega et al. and Wu could stem from the timing of their assessments, where self-perceptions were measured post-task.

This approach may have allowed students' awareness of their performance to bias their self-assessments. Additionally, both studies focused on complex linguistic tasks and concepts—such as sight translation, comprehensibility, and accentedness—that are not easily or consistently perceived by learners. For instance, some learners might associate comprehensibility with clear pronunciation, while others may link it to use of accurate and sophisticated vocabulary (Trofimovich et al., 2016).

On the other hand, our finding that learners' self-perceptions of their L2 abilities are not reliable predictors of their actual phonemic perception abilities is in line with a broader body of research indicating that language learners often struggle with accurately assessing their language skills. For example, Anderson (1982) found a nonsignificant negative correlation between ESL students' self-perceptions and their TOEFL scores, despite teachers' assessments positively correlating with the same scores. Likewise, Anderson noted a disconnect between students' self-assessments and their teachers' perceptions. Similarly, Blanche (1986) concluded that while self-assessments may reflect general language abilities, their reliability diminishes for specific competencies, as evidenced by the lack of significant correlation between students' self-assessments and their actual test performance. These findings support the conclusion that self-assessments are limited in their ability to predict actual linguistic abilities, particularly in specialized areas like phonemic perception.

Davidson and Henning (1985) also cautioned against placing too much trust in self-assessments, noting that students often overestimated their language abilities. Similarly, N. Ferguson (1978) reported a lack of significant correlation between self-assessments and scores on standardized language test among adult EFL learners while objective measures showed a correlation between standardized language test scores and the 'Speed of Identification of syntactic errors in written English.'

More recent studies, such as Masrai et al. (2022), identified vocabulary knowledge as the strongest predictor of academic success, followed by self-perception of language ability, while self-perceptions of L2 use were not significant predictors. They also reported significant positive correlations between Saudi EFL learners' actual and perceived vocabulary knowledge on one hand and self-perceptions of their abilities in listening, speaking, reading, and writing on the other hand. While Masrai et al. (2022) did not attempt to directly assess the correlation between measures of actual lexical knowledge and learners' self-perceptions of such knowledge, their findings are interesting. This is because they demonstrate that while objective measures of actual linguistic knowledge are significant predictors of academic success, learners' subjective self-perceptions are not. This further supports the idea that perceived linguistic competence cannot be reliably used as a valid measure of actual performance.

The critical gap between learners' self-assessments and their actual linguistic capabilities reflects a prevalent lack of self-confidence among participants, which is not uncommon among L2 learners (Clément et al., 1980; De Saint Léger, 2009; Léger & Storch, 2009; Metallidou & Efklides, 2001). Such underconfidence in one's linguistic abilities aligns with overall research findings that Arab EFL students are more likely to exhibit low self-confidence (Hamouda, 2013). This could be explained in the broader Arab cultural context, where an emphasis on modesty or a focus on group achievements over individual accomplishments may lead learners to undervalue their individual skills (Markus & Kitayama, 1991). In particular, this lack of confidence may contribute to heightened foreign language anxiety (Clément, 1987; Horwitz et al., 1986; MacIntyre, 2017; MacIntyre & Gardner, 1991; Young, 1991) and subsequently hinder language learning (Cheng, 2005; Horwitz et al., 1986; Krashen, 1982).

Perhaps the most interesting finding of our study is the persistent trend of cognitive bias consistent with the Dunning-Kruger effect (Dunning, 2011; Kruger & Dunning, 1999) While the overall correlation was weak, learners with lower actual phonemic perception abilities tended to overestimate their skills while higher-performing learners generally underestimated their performance. In this regard, our study joins existing studies (Blanche & Merino, 1989; Brantmeier & Vanderplank, 2008; N. Ferguson, 1978; Ortega et al., 2022; Saito et al., 2020; Suzuki, 2015; Trofimovich et al., 2016) in providing further empirical evidence of the Dunning-Kruger effect and extends such evidence to include Arab EFL learners' perceived phonemic abilities.

The tendency of higher-performing learners to underestimate their abilities and of lower-performing learners to overestimate their skills, as well as the large variability in both metacognitive awareness and actual linguistic performance—as indicated by the large score variance—underscores the complexity inherent in the study of learners' perceptions. In particular, this tendency suggests that even though they are objectively more skilled, more proficient learners may be more aware of the nuances and complexities of phonemic perception, leading them to be more self-critical and cautious in their assessments (Brantmeier & Vanderplank, 2008).

On the other hand, learners with lower perception skills may not yet have developed the necessary metacognitive strategies to accurately evaluate their phonemic abilities. Lacking these strategies, they may fail to recognize their limitations, leading to inflated self-assessments (Schraw & Dennison, 1994; Vandergrift & Goh, 2012; Wenden, 1998). An alternative plausible interpretation is that less proficient learners might still be in the early stages of interlanguage (Selinker, 1972) and are thus relying more heavily on transfer from their L1 (Lado, 1957) rather than noticing the salient features of the L2 phonemic system. This might have clouded their ability to realize gaps between their phonemic performance and the expected performance (R. Schmidt, 2001) and caused them to interpret L2 sounds based on the phonemic categories they had previously developed in their L1 (Aoyama et al., 2004; Flege et al., 1997).

Combined, our study findings highlight significant theoretical, epistemological, and pedagogical implications. From a theoretical perspective, the non-significant weak correlation between perceived and actual phonemic perception scores reveals a disconnect between learners' metacognitive awareness of their perceptual competence and their actual perceptual performance. While some SLA theories link high metacognitive awareness with L2 learning success, our results suggest that accurate metacognitive awareness is not a necessary condition for success. This suggests that individual differences play a significant role in the accuracy of L2 learner's self-perceptions of their

phonemic ability. Research has shown that individual differences in learners' cognitive abilities, age, native language, past experience can influence their self-perceptions (Blanche & Merino, 1989; Butler & Lee, 2010; Davidson & Henning, 1985; Janssen-van Dieten, 1989; Kathy Heilenman, 1990; Patri, 2002; Suzuki, 2015; Suzukida, 2024). Individual differences in affective factors such as motivation and anxiety can play a role, too (Cheng, 2005; Dörnyei & Skehan, 2003; Dörnyei & Ushioda, 2021; MacIntyre et al., 1997) .

From an epistemological perspective, our results illustrate the complexity of self-assessment among EFL learners, particularly in phonemic perception and reaffirm the subjective nature of self-perception. They underscore the legitimacy of our serious concerns about over-reliance on learner perceptions. Over-reliance on learner perceptions in L2 research often arises from the practicality and ease of administering self-assessment questionnaires. Studies frequently use self-perception as a proxy for actual performance due to its cost-effectiveness and time efficiency (Brantmeier & Vanderplank, 2008). However, the results of this study challenge the assumption that learners' self-reports can reliably replace objective performance measures and calls for a reevaluation of statements such as "self-ratings can be reliable and valid measures of communicative language abilities" (Bachman & Palmer, 1989, p. 22).

Bachman and Palmer's (1989) study has been cited more than 200 times as evidence "that self-ratings can be reliable and valid measures of communicative language abilities" (p. 22). Their research, which examined the construct validity of self-assessment, involved 116 non-native English speakers completing a 21-item multiple-choice self-rating test assessing language traits, difficulty experienced, and recognition of these traits in input. Bachman and Palmer's (1989) research has highlighted the nuanced relationships between test design and data accuracy and made significant contributions to the field. Key contributions include (a) their innovative use of a multitrait-multimethod (MTMM) design combined with confirmatory factor analysis (CFA) to assess the construct validity of self-assessment, (b) their identification of test design factors that enhance accuracy, such as the effectiveness of questions related to perceived difficulty, and (c) their demonstration that self-ratings' reliability varies depending on the language traits measured.

However, without measuring criterion-related validity, the assessment of how well a test correlates with an external measure of linguistic ability (Bachman, 1990; Mills & Jordan, 2022), it remains unclear whether the internal consistency of self-assessments alone translates into accurate self-perceptions of linguistic competence. Exploring such criterion-related evidence by holding a direct comparison between self-perceptions and an external benchmark of linguistic ability, our study reveals that self-perceptions are not necessarily valid measures of actual language abilities. As a result, ESL learners' self-perceptions should be treated with caution, as they may not provide a trustworthy measure of learner linguistic performance. Whenever possible, future research studies should supplement these subjective measures by objective assessments to ensure a more balanced and accurate understanding of the variables under investigation.

From a pedagogical perspective, our findings highlight the need for targeted educational interventions to help learners improve their phonemic perception accuracy. While the learners in our study performed above average, there remains room for improvement. Research demonstrates that L2 learners can better discriminate phonemes through High Variability Phonetic Training, a perception-based training approach, (Thomson, 2011; Wang & Munro, 2004; Wong, 2015) and related perceptual training activities, such as phoneme discrimination and identification (Bradlow et al., 1999; Thomson, 2011).

Our findings also suggest that educational strategies need to address the discrepancies between perceived and actual phonemic perceptions. When learners underestimate their abilities, they may be less likely to engage in tasks that challenge their phonemic perception skills, which could hinder their development. Research emphasizes that learners who lack confidence in their language abilities are less likely to take risks in language use, which can limit opportunities for growth and improvement (Dörnyei & Ushioda, 2021). Teachers can mitigate this issue by incorporating metacognitive training into phonemic perception instruction. This may entail instructing learners to actively listen for specific phonemic contrasts in spoken language, compare their own pronunciation to that of native speakers, and seek feedback on their performance. Teachers should also prioritize providing learners with regular and specific feedback on their phonemic perception skills.

While the findings suggest that learners are not always accurate in their self-assessments, encouraging learners to engage in self-reflection and self-assessment tasks, with the support of clear criteria and guided feedback, may help them self-regulate their learning (Dann, 2002; Oscarson, 1989, 1997; Paris & Paris, 2001) and develop more accurate self-assessments over time (Boud, 1995; Butler & Lee, 2010).

Finally, the presence of the Dunning-Kruger effect in this study suggests that learners at different levels of proficiency may require different types of support. For lower-performing learners, interventions should focus on increasing awareness of phonemic distinctions and helping learners to develop strategies for improving their perception abilities. For higher-performing learners, teachers should focus on building confidence and encouraging learners to continue developing their skills through more advanced tasks. Computer assisted pronunciation training can be effectively used to tailor training and feedback to individual learner needs (Thomson, 2011).

7. Limitations

While this study provides valuable insights, several limitations should be acknowledged. First, the study's reliance on a relatively small, homogeneous sample of 58 male Arabic-speaking EFL learners limits the generalizability of its findings. The participants were all from a specific linguistic and cultural background (Qassimi Arabic speakers) and enrolled in the same university, which may not reflect the diversity of the broader EFL learner population. Future studies should include a more diverse participant pool in terms of gender, regional dialects, and educational contexts to enhance external validity.

Second, the study's focus on English vowel perception, while insightful, may not fully capture the complexity of phonemic perception across different linguistic structures. Further research should examine other phonemic categories, such as consonants or suprasegmental features (e.g.,

intonation), to provide a more holistic view of learners' phonological difficulties in EFL contexts. Addressing these limitations in future research will help strengthen the conclusions drawn and support the development of more targeted pedagogical interventions.

8. Conclusion

In conclusion, this study provides valuable insights into the relationship between perceived and actual phonemic perception among Arabic-speaking English as a Foreign Language (EFL) learners. The findings indicate a significant gap between learners' self-assessments and their actual abilities, with a majority of participants underestimating their phonemic perception skills. This discrepancy aligns with the Dunning-Kruger effect, where lower-performing learners overestimate their abilities, while higher-performing learners tend to underestimate theirs. The weak correlation between perceived and actual phonemic perception abilities further challenges the validity of relying on self-assessments as an accurate measure of linguistic competence. These results highlight the need for more objective, performance-based measures in language learning research and pedagogy. Additionally, the study underscores the importance of targeted educational interventions, such as metacognitive training, to help learners develop more accurate self-assessment skills and enhance their phonemic perception abilities. Importantly, this research has implications beyond the specific context of Arabic-speaking EFL learners. The insights gained can resonate with educators and researchers globally, as they address a common challenge in language acquisition across diverse linguistic backgrounds. The findings can inform educators and policymakers worldwide about the discrepancies between perceived and actual abilities, emphasizing the need for improved assessment methods in diverse educational contexts. Future research should continue exploring the role of individual differences and cognitive biases in language learning to create more effective instructional strategies. Overall, this study contributes to a deeper understanding of the complexities of learner perception and its implications for second language acquisition.

Acknowledgments

Not applicable

Authors' contributions

FA: Writing– original draft, Writing – review & editing, Conceptualization, and Methodology; SI: Writing– original draft, Writing – review & editing, Conceptualization, and formal Analysis. Authors contributed equally to the study and share first authorship.

Funding

The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024)

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Anderson, P. L. (1982). Self-esteem in the foreign language: A preliminary investigation. *Foreign Language Annals*, *15*(2), 109–114. https://doi.org/10.1111/j.1944-9720.1982.tb00234.x

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250. https://doi.org/10.1016/S0095-4470(03)00036-6

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing*, *6*(1), 14–29. https://doi.org/10.1177/026553228900600104

Bercher, D. A. (2012). Self-monitoring tools and student academic success: When perception matches reality. *Journal of College Science Teaching*, *41*(5), 26–32. https://doi.org/10.2505/3/jcst12_041_05

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

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Flege* (pp. 13–34). John Benjamins. https://doi.org/10.1075/lllt.17.07bes

Blanche, P. (1986). *The relationships between self-assessments and other measures of proficiency in the case of adult foreign language learners* [Master's Thesis]. University of California, Davis.

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language Learning*, *39*(3), 313–338. https://doi.org/10.1111/j.1467-1770.1989.tb00595.x

Boud, D. (1995). *Enhancing learning through self assessment*. Kogan Page.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977–985. https://doi.org/10.3758/BF03206911

Brantmeier, C., & Vanderplank, R. (2008). Descriptive and criterion-referenced self-assessment with L2 readers. *System*, *36*(3), 456–477. https://doi.org/10.1016/j.system.2008.03.001

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology*, *64*(1), 74–95. https://doi.org/10.1080/17470218.2010.499174

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology, Vol. 3: Cognitive development* (4th ed., Vol. 3, p. 77—166). Wiley.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, *27*(1), 5–31. https://doi.org/10.1177/0265532209346370

Cheng, Y. (2005). EFL learners' listening comprehension anxiety. *English Teaching and Learning*, *23*(9), 25–44. https://www.researchgate.net/publication/299747187_EFL_Learners'_Listening_Comprehension_Anxiety

Chomsky, N. (1966). *Topics in the theory of generative grammar*. Mouton.

Clément, R. (1987). Second language proficiency and acculturation: An investigation of the effects of language status and individual characteristics. *Journal of Language and Social Psychology*, *5*(4), 271–290. https://doi.org/10.1177/0261927X8600500403

Clément, R., Gardner, R. C., & Smythe, P. C. (1980). Social and individual factors in second language acquisition. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *12*(4), 293–302. https://doi.org/10.1037/h0081081

Dann, R. (2002). *Promoting assessment as learning: Improving the learning process*. Routledge.

Davidson, F., & Henning, G. (1985). A self-rating scale of English difficulty: Rasch scalar analysis of items and rating categories. *Language Testing*, *2*(2), 164–179. https://doi.org/10.1177/026553228500200205

Davies, A. (2003). *The native speaker: Myth and reality.* Multilingual Matters.

De Saint Léger, D. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals*, *42*(1), 158–178. https://doi.org/10.1111/j.1944-9720.2009.01013.x

Dekeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 125–151). Cambridge University Press. https://doi.org/10.1017/CBO9781139524780.007

Dörnyei, Z., & Dewaele, J.-M. (2023). *Questionnaires in second language research: Construction, administration, and processing* (3rd ed.). Routledge.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Blackwell. https://doi.org/10.1002/9780470756492.ch18

Dörnyei, Z., & Ushioda, E. (2021). *Teaching and researching motivation* (3rd ed.). Routledge. https://doi.org/10.4324/9781351006743

Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 44, pp. 247–296). Elsevier. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.

Evans, B. G., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics*, *68*, 15–31. https://doi.org/10.1016/j.wocn.2018.01.002

Ferguson, C. A. (1959). Diglossia. *Word, 15*, 325–340. https://doi.org/10.1080/00437956.1959.11659702

Ferguson, N. (1978). Self-assessment of listening comprehension. *International Review of Applied Linguistics*, *16*, 149–156. https://doi.org/10.1515/iral.1978.16.1-4.149

Field, J. (2009). *Listening in the language classroom*. Cambridge University Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

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. 233–277). York Press.

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. https://doi.org/10.1006/jpho.1997.0052

Gass, S. M., Behney, J., & Plonsky, L. (2020). *Second language acquisition: An introductory course* (5th ed.). Routledge.

Goh, C. (1997). Metacognitive awareness and second language listeners. *ELT Journal, 51*(4), 361–369. https://doi.org/10.1093/elt/51.4.361

Goh, C. C. M., & Vandergrift, L. (2022). *Teaching and learning second language listening: Metacognition in action* (2nd ed.). Routledge.

Hamouda, A. (2013). An exploration of causes of Saudi students' reluctance to participate in the English language classroom. *International Journal of English Language Education, 1*(1), 17–34. https://doi.org/10.5296/ijele.v1i1.2652

Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties.* (p. 419). Georgetown University Press. https://doi.org/10.1353/lan.2007.0094

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://www.jstor.org/stable/327317

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning*, *4*(2), 161–176. https://doi.org/10.1007/s11409-009-9042-8

Hymes, D. (1966). Two types of linguistic relativity (with examples from Amerindian ethnography). In W. Bright (Ed.), *Sociolinguistics* (pp. 114–167). De Gruyter Mouton. https://doi.org/10.1515/9783110856507-009

Ingham, B. (1994). *Najdi Arabic*. John Benjamins.

Janssen-van Dieten, A. M. (1989). The development of a test of Dutch as a second language: The validity of self-assessment by inexperienced subjects. *Language Testing*, *6*(1), 30–46. https://doi.org/10.1177/026553228900600105

Kathy Heilenman, L. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing*, *7*(2), 174–201. https://doi.org/10.1177/026553229000700204

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.

Léger, D. de S., & Storch, N. (2009). Learners' perceptions and attitudes: Implications for willingness to communicate in an L2 classroom. *System*, *37*(2), 269–285. https://doi.org/10.1016/j.system.2009.01.001

Lightbown, P., & Spada, N. (2021). *How languages are learned* (5th ed.). Oxford University Press.

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New Insights Into Language Anxiety* (pp. 11–30). Multilingual Matters. https://doi.org/10.21832/9781783097722-003

MacIntyre, P. D., & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning*, *41*(1), 85–117. https://doi.org/10.1111/j.1467-1770.1991.tb00677.x

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety.

Language Learning, *47*(2), 265–287. https://doi.org/10.1111/0023-8333.81997008

Mackey, A., & Gass, S. M. (2022). *Second language research: Methodology and design* (3rd ed.). Routledge. https://doi.org/10.4324/9781003188414

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Masrai, A., El-Dakhs, D. A. S., & Yahya, N. (2022). What predicts academic achievement in EMI courses? Focus on vocabulary knowledge and self-perceptions of L2 skills. *SAGE Open*, *12*(2), 1–13. https://doi.org/10.1177/21582440221101044

McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 213–228). Academic Press. https://doi.org/10.1016/B978-012589042-7/50009-8

Metallidou, P., & Efklides, A. (2001). The effects of general success-related beliefs and specific metacognitive experiences on causal attributions. In *Trends and prospects in motivation research* (pp. 325–347). Kluwer Academic Publishers.

Mills, G. E., & Jordan, A. W. (2022). *Educational research: Competencies for analysis and applications* (13th ed.). Pearson.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press.

Ortega, M., Mora, J. C., & Mora-Plaza, I. (2022). L2 learners' self-assessment of comprehensibility and accentedness: Over/under-estimation, effects of rating peers, and attention to speech features. In J. Levis & A. Guskaroska (Eds.), *Proceedings of the 12th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–13). Iowa State University Digital Press. https://doi.org/10.31274/psllt.13354.

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, *6*(1), 1–13. https://doi.org/10.1177/026553228900600103

Oscarson, M. (1997). Self-assessment of foreign and second language proficiency. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: Volume 7: Language testing and assessment* (pp. 175–187). Kluwer Academic Publishers.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Heinle & Heinle.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4

Patri, M. (2002). The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing*, *19*(2), 109–131. https://doi.org/10.1191/0265532202lt224oa

Roach, P. (2009). *English Phonetics and Phonology*. Cambridge University Press.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, *15*(1), 1–20. https://doi.org/10.1177/026553229801500101

Ryding, K. (2005). *A reference grammar of Modern Standard Arabic*. Cambridge University Press.

Saito, K., Trofimovich, P., Abe, M., & In'nami, Y. (2020). Dunning-Kruger effect in second language speech learning: How does self perception align with other perception over time? *Learning and Individual Differences*, *79*, 1–10. https://doi.org/10.1016/j.lindif.2020.101849

Saville-Troike, M. (2016). *Introducing second language acquisition* (3rd ed.). Cambridge University Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press. https://doi.org/10.1017/CBO9781139524780.003

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351–371. https://doi.org/10.1007/BF02212307

Schulz, R. A. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar instruction and corrective feedback: USA-Colombia. *The Modern Language Journal*, *85*(2), 244–258. https://doi.org/10.1111/0026-7902.00107

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(3), 209–231. https://doi.org/10.1515/iral.1972.10.1-4.209

Stan, I. (2024). Self-assessment versus objective proficiency evaluation in English as a foreign language among Italian first-year university students. *Open Journal of Modern Linguistics*, *14*(04), 543–558. https://doi.org/10.4236/ojml.2024.144028

Suzuki, Y. (2015). Self-assessment of Japanese as a second language: The role of experiences in the naturalistic acquisition. *Language Testing*, *32*(1), 63–81. https://doi.org/10.1177/0265532214541885

Suzukida, Y. (2024). Delving into L2 learners' perspective: Exploring the role of individual differences in self-evaluation of L2 speech learning. *Languages*, *9*(3), 1–26. https://doi.org/10.3390/languages9030109

Thomson, R. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal*, *28*, 744–765. https://doi.org/10.11139/cj.28.3.744-765

Trofimovich, P., Isaacs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: Investigating self- and other-perception of second language speech. *Bilingualism: Language and Cognition*, *19*(1), 122–140. https://doi.org/10.1017/S1366728914000832

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539–552. https://doi.org/10.1016/j.system.2004.09.011

Watson, J. C. E. (2007). *The phonology and morphology of Arabic*. Oxford, England: Oxford University press.

Wenden, A. L. (1998). Metacognitive knowledge and language learning. *Applied Linguistics*, *19*(4), 515–537. https://doi.org/10.1093/applin/19.4.515

Wesely, P. M. (2012). Learner attitudes, perceptions, and beliefs in language learning. *Foreign Language Annals*, *45*(1), 98–118. https://doi.org/10.1111/j.1944-9720.2012.01181.x

Wong, J. W. S. (2015). The impact of L2 proficiency on vowel training. In J. A. Mompean & J. Fouz-Gonz á lez (Eds.), *Investigating English pronunciation* (pp. 219–239). Palgrave Macmillan UK. https://doi.org/10.1057/9781137509437_10

Wu, Z. (2019). Text characteristics, perceived difficulty and task performance in sight translation: An exploratory study of university-level students. *Interpreting*, *21*(2), 196–219.

Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, *75*(4), 426–439. https://doi.org/10.2307/329492