International Journal of Education, Culture and Society
2022; 7(1): 33-41
http://www.sciencepublishinggroup.com/j/ijecs
doi: 10.11648/j.ijecs.20220701.16
ISSN: 2575-3460 (Print); ISSN: 2575-3363 (Online)

Learning Proper VOT of English Velar Stops by Kuwaiti Undergraduate Students from VOT and Optimality Theory Perspectives

Sultan Melfi Aldaihani*, Shorouq Khalid Al-Houti, Yousef Masad Alenezi

The English Department, College of Basic Education, Al-Ardhiya, Kuwait

Email address:
dr.aldaihani@hotmail.com (S. M. Aldaihani), dr.alhouti@gmail.com (S. K. Al-Houti), englishq8@hotmail.com (Y. M. Alenezi)
*Corresponding author

To cite this article:
Sultan Melfi Aldaihani, Shorouq Khalid Al-Houti, Yousef Masad Alenezi. Learning Proper VOT of English Velar Stops by Kuwaiti Undergraduate Students from VOT and Optimality Theory Perspectives. *International Journal of Education, Culture and Society*. Vol. 7, No. 1, 2022, pp. 33-41. doi: 10.11648/j.ijecs.20220701.16

Received: January 2, 2022; **Accepted**: February 6, 2022; **Published**: February 16, 2022

Abstract: An important issue in the recent L2 literature is whether adult learners can perceive a gradient phonetic difference between L1 and the corresponding L2 phonemes. This leads us to further judge whether such learners can acquire allophonic contrast in L2. The same issue has been addressed in the current study with reference to Kuwaiti adult learners of English. The aim of this study is to analyze velar stops /k, g/ as perceived and produced by Kuwaiti Arabic (KA) learners of English. In this study, comprehension is measured by a perception test, and pronunciation is measured by obtaining readings of voice onset time (VOT) in Praat software. The velar stops will also be analyzed using the framework of the classic version of optimality theory (OT). OT became one of the major generative frameworks in the field of generative phonology. OT is a constraint-based theory of phonology advanced by Prince and Smolensky. In the perception test, participants were asked to listen to words containing target sounds, which were "keys," "skis," and "geese," and note on paper which English word they had heard. In the discrimination task, these words were presented in pairs, and the participants were asked to say whether they had heard a single word twice or two words together. In the production test, they read from a list of words, including "keys," "skis," and "geese," and their productions were recorded. A large group of Kuwaiti English learners perceived and produced these velar stops of English. Their discrimination of voiced and voiceless velar stops was excellent, but their identification of [g] was weak. Their identification of the voiceless velar stop /k/ was also native-like. In production, they were native-like in aspirated [k^h]. In the production of the unaspirated voiceless velar stop [k], they were not native-like, but they had developed an understanding of this allophone of English. In [g], some students were native-like, some were still learning, and some had only relied on L1 transfer.

Keywords: Second Language, Phonology, Perception, Production, Optimality Theory, English, Kuwaiti Arabic

1. Introduction

English and Kuwaiti Arabic (KA) have different phonemic inventories in vowels and consonants [3]. Some consonants of one language do not exist in the other. For example, Arabic has two laryngeal fricatives [3], but English has one [38]. In some examples, both languages have similar consonants in their phonemic inventories but not the same because their corresponding sounds are slightly different from each other. Such sounds with slight phonetic differences are more difficult to learn than different sounds [10]. One example of these is velar plosives /k, g/ in the English and KA languages. First, they are similar but have phonetic differences. English has aspiration at phonetic allophonic level, which KA does not have. Second, English voiced stops are phonetically voiceless but KA voiced stops are pre-voiced. These minor phonetic differences may give a KA learner of English some trouble. There is no previous research on this topic with reference to the problems faced by KA learners of English. Further, there is no research carried out in KA learners of English velar stops from the optimality theory (OT) perspective. This study fills these gaps and studies KA learners' acquisition of velar stops in English L2 with an OT framework [36].

2. Literature Review

2.1. Voice Onset Time (VOT) for Stops

This paper examines the comprehension and pronunciation of English velar stops /k, g/ by Kuwaiti learners of English. Comprehension is measured by a perception test, and pronunciation is measured by obtaining readings of voice onset time (VOT) in Praat software [7]. The VOT for plosives is the most commonly used acoustic correlate to determine whether the pronunciation of language learners is correct or incorrect. The concept of VOT was first proposed by [23].

According to Lisker, L et al, VOT is the measurement of the duration of time from the burst of a plosive until the beginning of a vocal fold vibration for the creation of voicing for the following vowel in a syllable [23]. The VOT of a stop can only be measured if it is followed by a vowel. There are three possible types of VOT. If the vocal folds of a speaker start vibrating before the active and passive articular of the speaker separate for the burst phase, such a VOT is called pre-voicing, and such consonants are called truly voiced. The voicing duration of a truly voiced stop is always mentioned in negative values. If the VOT is positive but small, the produced stop will be declared an unaspirated stop or a stop with a short lag VOT (/k/ in skis), and if it is a longer VOT, then it will be considered an aspirated stop or stop with long lag VOT (/k/ in keys). English has a shorter and longer VOT in stops, but Arabic dialects have negative and positive VOTs for voiced and voiceless stops, respectively. This context creates multiple difficulties for Arab learners in learning the correct pronunciation of English. Another difference is that English differentiates between short and long VOT, but Arabic dialects have only one range of VOT without any phonemic or allophonic variance in VOT. These differences may cause big trouble for Arab learners in learning the plosives of English. Many studies have been conducted to highlight this problem in the Arab world, but none has focused on Kuwaiti learners of English. This study fills this gap and can solve the problems of learning English velar stops faced by Kuwaiti English learners.

2.2. Optimality Theory (OT)

At the beginning of the 1990s, OT became one of the predominant generative frameworks in the area of generative phonology. OT is a constraint-based theory of phonology, advanced by Prince, A et al. [36]. The key to its success is its applicability in all fields of grammar, providing a new perspective on a wide range of problems in linguistics and providing solutions to problems that were not treatable by old theories (see [9]). For example, the problems with conspiracy [20] and opacity were solved by OT (see Stratal OT; [6, 19]).

The major shift OT brought about in phonology is from a rule-based model to an output-based one, a move that was foreshadowed in several other publications [36]. OT has played a vital role in generative phonology and has helped improve our understanding of how grammar works in general. It is intended to be a framework that captures cross-linguistic generalizations.

Cross-linguistically, Prince and Smolensky took the lead in explaining the OT approach to the different types of phonological phenomena [36]. Subsequently, several modifications have been proposed [36, 30, 31, 32, 14]. Various OT models have offered comprehensive solutions to different phonological phenomena, such as Stratal OT [6, 19], correspondence theory [32], and harmonic serialism [36, 26]. In this study, classic OT will be used to provide phonological evidence on KA learners' ability with English velar stops.

OT is an analytical framework that allows the exploration of both the formal properties of grammar and its developmental and cross-linguistic variations. OT is a phonological formalism in which the pronounced form of an utterance (the output) is chosen from among other possible candidates that are all simultaneously evaluated and directly compared to the lexical representation (the input) by a ranked set of violable and universal constraints. In other words, OT suggests an output-based grammar where constraints replace rules and processes. These constraints are universal, violable, and ranked, and their objective is to match output with input in the most harmonic form possible. Constraints account for both universal generalizations and language-specific forms. This means that universal grammar provides both the range of probable outputs and a set of constraints to confirm that the selected outputs are optimal for the spoken language.

The universal set of OT constraints contains three constraints, faithfulness, markedness, and alignment, which are violable and able to form the core constraints in OT. A faithfulness constraint is needed so the mental representation (input) and the surface representation (output) can be identical in terms of their phonological characteristics, meaning that outputs must preserve the properties of their basic lexical forms to have some kind of similarity between the output and its input [14]. This means that faithfulness constraints require input and output forms to be identical. If segments are added, omitted, or undergo featural changes, FAITH is violated. Faithfulness constraints are composed of DEP, a constraint against insertion; MAX, a constraint against the deletion of segments; and IDENT, a constraint against the changes in place or manner of segment production. DEP, IDENT, and MAX are good examples of faithfulness constraints:

The definition of constraint "DEP": every segment of S2 has a correspondent segment in S1.

Range (*R*) = S2 [14].

*The definition of constraint "IDENT [F]: c*orresponding segments must be identical in features [14].

*The definition of constraint "MAX": e*very segment of S1 has a correspondent segment in S2.

Domain (*R*) = S1 [14].

The concept of markedness means that all kinds of linguistic structures have two different values: one is marked structure, and the other is unmarked. A marked value is one that is avoided, whereas an unmarked value is one that is cross-linguistically preferred [14]. Markedness constraints make direct statements about the marked or unmarked configurations of a surface form. For example, markedness constraints either prevent marked configurations (such as *COMPLEX and NOCODA) or demand unmarked characterization (such as PEAK and ONSET).

Examples of markedness constraints

The definition of constraint " COMPLEX"* is that *n*o more than one C or V may be associated with any syllable node [14].

The definition of constraint "NO CODA": Syllables must not have codas [14].

The definition of constraint "PEAK": a syllable must have a vocalic peak [14].

The definition of constraint "ONSET" is that *s*yllables must have onsets [14].

In other words, markedness constraints enforce well-formedness and require outputs to be unmarked or simplified in structure and/or segments. These include the following: *COMPLEX ONSET, a constraint against consonant cluster syllables initially; *CODA, a constraint against final consonants; *LIQUIDS, a constraint against liquids; and *FRICATIVES, a constraint against fricatives. This shows that the two main functions of OT are to generate (GEN) and evaluate (EVAL) an output. Every language uses the same set of constraints but prioritizes them differently. When two different syllabifications are compared, conflicts due to ranking constraints will occur. For example, the CVCCV input of English ranks NoCoda>>*Complex Onset as in the output CV.CCV (re.ply), whereas Cantonese ranks *Complex onset>>NoCoda as in the output CVC.CV (yap.le) [36].

Markedness constraints penalize disfavored outputs (e.g., NoCoda), and faithfulness constraints penalize changes to input (e.g., NoDeletion), meaning that faithfulness constraints interact with markedness constraints to protect the underlying form from any unpredicted phonological alterations. These constraints are universal, but every language has its own specific rankings. If the surface form of a specific language accepts certain phonological changes such as assimilation, pharyngealization, and insertion, the markedness constraint will then be higher ranked than the faithfulness constraint.

According to McCarthy, J et al, generalized alignment (GA) is considered a sub-theory of OT that accounts for footing and extends to advance the correspondence relationship between prosody and morphology rather than considering the correspondents as templatic, MCat = PCat [31], as in previous work [28]. Such constraints demand the alignment of edges between phonological and/or morphological constituents, meaning that the right or left edge of one prosodic or grammatical category coincides with the right or left edge of another. Alignment constraints are normally used in OT analyses, especially within prosodic morphology. The following constraint is a good example of GA:

The definition of constraint "ALIGN (FT R, PRWD R)":

The right edge of every foot must align with the right edge of the prosodic word [31].

2.2.1. Architecture of OT

OT is an analytical framework of how the input–output relationship is governed by well-formedness constraints. In OT, the pronounced form of an utterance (the output) is chosen from among other possible candidates that are all simultaneously evaluated and directly compared to the lexical representation (the input) by a ranked number of violable and universal constraints [36, 30, 31, 14].

Classic OT is limited to a single derivation, producing only the ultimate output through GEN and EVAL. It has an unrestricted Gen, which respects *inclusiveness and freedom of analysis* because all generated ultimate outputs are recognized at once; therefore, some candidates will show the influence of different phonological processes simultaneously (see [28]). More precisely, the outputs of the OT's GEN allow unlimited changes to the input in one step.

The following diagrams show classic OT, with its single operation through GEN and EVAL. A ranking of universal constraints chooses the most optimal candidate from this unlimited candidate set as output until there is convergence. When the output is identical to the old input, the derivation has converged, and so it ends. However, for Classic OT, EVAL selects the most optimal candidate through a language-specific hierarchy of ranked constraints on the first pass, as shown below.

Graphic representation of classic OT structure

Input→ GEN→ Candidate-set → EVAL→ Output → (unrestricted GEN) [36].

2.2.2. Previous OT Studies

The analysis of a phonological process is based mainly on the principles of generative phonology. The term "phonological process" is used here to express the way in which certain sound segments in specific environments undergo phonological changes or alternations. The phonological processes include resyllabification, insertion, pharyngealization, vowel harmony, assimilation, etc. Many studies have been conducted to analyze different phonological processes in both different Arabic dialects and other languages. Few studies of Arabic dialects have treated such segmental changes from the OT perspective.

Few studies have been carried out on Arabic dialects with a framework of OT ([25, 27, 32, 39, 14] among others). In addition, Mustafawi tackled the variable consonantal alternation in Qatari Arabic from the OT perspective [33], whereas Al-Mohanna investigated syllabification and metrification in Urban Hijazi Arabic [4]. Two studies on Syrian Arabic were conducted by Adra and Al-Omar, tackling epenthesis, syncope, stress [2, 5], and emphasis spreading (ES) with the OT framework [2]. Another study on opacity and transparency in the phonology of Makkan Arabic focusing on insertion, stress, and syncope used Stratal OT (see [13]), and Rakhieh used Stratal OT as a framework to study insertion, stress, and deletion in Ma'ani Arabic/Jordanian dialect [37]. To the best of my knowledge, there has been only one attempt to analyze KA within the framework of OT. Aldaihani investigated assimilation, pharyngealization, and insertion within the framework of harmonic serialism and OT [3]. By contrast, there is no previous research on KA learners' acquisition of velar stops /k, g/ in English L2 using optimality theory. This encourages the researcher to be one of the first researchers to scrutinize and explain the effect and impact of classic OT formalism on treating the English velar stops learned by KA learners of English.

3. Research Methodology

In this study, we will analyze our findings using Praat [7] and use classic OT [36] to formalize generalizations and provide phonological evidence for the acquisition of English velar stops by KA learners. To determine the reliability level of these data, a Cronbach's alpha reliability test was applied to the repetitions. There are many versions of OT, but classic OT will be the selected framework for treating the data in this study.

To get a clear concept of the perception and production of velar stops by Kuwaiti learners of English, we arranged two perception and one production test with 106 Kuwaiti learners of English in the College of Basic Education in Kuwait (English Department). All the participants are women. In the perception test, they were asked to listen to words containing target syllables, which were "keys," "skis," and "geese," and write down on paper which English word they had heard. In a second discrimination task, these words were presented in pairs, and they were asked to decide whether they had heard a single word twice or two words together.

In the production test, they read words from paper including "keys," "ski," and "geese." The performances were recorded. The VOTs of the word-initial velar stops were taken using standard methods as mentioned in [16] using Praat [7]. There were three repetitions of each word. A Cronbach's alpha reliability test was applied to the repetitions to determine the reliability level of this data. The results of the reliability test are given below:

Table 1. *Reliability Test Results on Kuwaiti Learners' Productions.*

S. No.	Consonant	Cronbach's alpha	% reliability
1	[k^h]	.851	85%
2	[k]	.837	84%
3	[g]	.382	38%

The reliability coefficient was excellent in both aspirated and unaspirated voiceless velar stops. Normally, a Cronbach's alpha value of .7 or above is considered excellent reliability in research studies ([22, 40]). It was below the standard level for [g], but this does not mean the reliability is low. As the data will show in the later sections, some of the participants produced [g] with negative VOT and others with positive VOT. Because of this, the reliability coefficient went lower, which is an artefact of statistics but not an indicator of low reliability.

Two other groups of native speakers of English and KA were also recorded for L1 VOTs of English and KA. A total of fifty-three monolingual native speakers of KA produced the written Arabic words *keef* ("mood") and *gasi* ("tough"), which were recorded for obtaining VOT values. A Cronbach's alpha reliability test applied on these data yielded above 90% reliability in these repetitions. The results of the test are written below:

Table 2. *Reliability Test Results on Native KA Speakers' Productions.*

S. No.	Consonant	Cronbach's alpha	% reliability
1	[k]	.901	90%
2	[g]	.943	94%

On the same pattern, 11 native speakers of English were requested to speak the words "keys," "skis," and "geese," which were recoded with repetitions. But this was applied to the VOTs of native speakers of English, and the results are given below:

Table 3. *Reliability Test Results on Native English Speakers' Productions.*

S. No.	Consonant	Cronbach's alpha	% reliability
1	[k^h]	.904	90%
2	[k]	.843	84%
3	[g]	.437	43%

The reliability of native speakers' production is also excellent, except with [g], the reason for which is the same as already discussed. Native English-speaking participants in England were paid for their time. L2 students in Kuwait were compensated with credit marks for their terminal assignments. The results are described in the following section. KA monolingual participants provided their services voluntarily.

4. Results

4.1. Perception Test Results

In the identification test, learners were asked to listen to and identify English words. They wrote their answers on paper. There were three repetitions of each target token. One mark was awarded for a correct identification. The results are averaged and summarized in the following table:

Table 4. *Identification Test Result.*

Words	Mean	Std. Deviation	Percentage
Geese	.87	1.10	28.85
Keys	2.9	.234	98.08
Skis	2.88	.58	96.15

These are the marks obtained by participants on an average of three marks. The difference between learners' performance on the identification of "geese" was significantly different from "keys" (t=19.046, p= .0001) and "skis" (t=16.422, p= .0001). However, the results for "keys" and "skis" were not significantly different (p= .357). The results indicate that Kuwaiti learners can identify both allophones of English /k/ equally well, but they cannot identify /g/ consonant produced by native speakers of English. a discrimination test was arranged for further confirmation of Kuwaiti learners' perception. In this test, each participant listened to each sound pair three times. They had to determine if they had heard two different words or a single word twice. One mark was awarded for each correct answer. The results of this test are found below:

Table 5. *Discrimination Test Results.*

Pairs	Mean	Std. Deviation	Percentage
Keys-Geese	2.58	.77	86
Geese-Geese	2.87	.40	95.67
Geese-Keys	2.42	.93	80.67

This result shows that the participants were very proficient in the discrimination between words starting with /k/ and /g/. These results show that Kuwaiti learners of English can discriminate between English /k/ from English /g/. But in the previous results, we noticed they could not identify English /g/ properly. These results are further analyzed in the analysis and discussion section. Production test results are given in the following subsection.

4.2. Production Test Results

To get an idea of native speaker VOT ranges, we recorded two words of KA, "keef" and "gasi," produced by 53 monolingual native speakers of KA, and we obtained VOT values for word-initial velar stops in these productions. The results are in the table below:

Table 6. *Native VOT Ranges for /k/ and / g/ in KA.*

S. No.	Words	Mean	Std. Deviation
1	Keef	57.25	13.058
2	Gasi	-68.36	19.62

These results prove that the KA voiceless velar /k/ is produced with aspiration in the range of a 57.25 ms VOT, but /g/ is produced with an almost 68.36 ms pre-voicing duration. These findings are in line with previous research that claims that Arabic voiceless stops are aspirated but voiced stops are truly voiced or pre-voiced [15, 18, 17, 1]. This confirms that KA is a voicing language according to the division of languages [12, 11]. Other voicing languages that are like Arabic include Japanese [34], Saraiki [43], Dutch [42], and Korean [41].

To get L1 VOT ranges of English stops, we also recorded 11 monolingual native speakers of British English and obtained VOT values of their velar stops. The results are given in the table below:

Table 7. *Native English Speakers' VOT Values for Velar Stops.*

	Mini mum	Maxi mum	Mean	Std. Deviation
Keys	67.67	114.00	89.42	14.41
Skis	10.33	61.33	27.24	14.38
Geese	13.67	61.33	31.30	12.31

These results confirm that the VOTs of aspirated velar stops in English are around 89.42 ms, but the VOTs of unaspirated velar stops are 27.24 ms and those of voiced velar stops are 31.30 ms. There was no significant difference between the VOTs of unaspirated velar stops and voiced velar stops in the speech of English native speakers (t=12.533, p= .0001). In other words, English native speakers produce voiced and voiceless unaspirated velar stops with the same VOTs. These findings are in accordance with the previous results taken in [8].

A total of fifty-three Kuwaiti learners of English produced English words starting with velar stops. Each target word had two repetitions, and the repetitions were averaged. The following table shows the results of the averaged VOT calculations.

Table 8. *VOTs of English Velar Stops by Kuwaiti Learners.*

	N	Mini mum	Maxi mum	Mean	Std. Deviation
Keys	106	46.00	130.33	85.15	20.19
Ski	106	34.33	125.67	66.34	19.94

These VOT values show that Kuwaiti students have developed two different VOT ranges for English /k/ in aspirated "keys" and unaspirated "skis" because the two VOTs differ significantly (t=10.182, p= .0001). The VOTs for the voiced velar stop /g/ were widely dispersed. Some participants produced this consonant with positive VOT and some with negative VOT. These values were considered separately.

We had 318 (106*3=318) productions of the word "geese" by Kuwaiti student learners of English. Among those, 198 were produced with positive VOT, and the mean of those productions was around 29 ms. By contrast, 90 productions had pre-voicing, and the mean VOT was -74 ms. The remaining 30 productions were with zero VOT. These results are analyzed in the following sections.

5. Analysis and Discussion

5.1. General Analysis and Discussion

In the perception test, we realized that Kuwaiti students are native-like at discrimination but weak at the identification of the voiced stop /g/. This is understandable. Discrimination is easier for anyone because one can just hear and feel the existing difference between the two different sounds, although s/he does not identify those consonants. But identification is only possible when we really understand the actual nature of a sound. Therefore, we see successful results of KA learners in the discrimination test and weak results in the identification test. When the nature of the errors of Kuwaiti students were analyzed, we came to know that all those who erred in the identification of /g/ had actually identified /g/ as [k]. This is also logical because, in their language, a voiced velar stop is one that is produced with negative VOT. But the /g/ sound used in the token for identification was produced by an English native speaker who produced this sound with positive VOT. As our statistics show, native speakers of English produced English /g/ with 31.30 ms VOT. For Kuwaiti learners, anything produced with this VOT range after the burst of a stop can only be a voiceless stop. We can conclude that learners' perceptions are influenced by their L1. The weak result in the identification of /g/ is because of the negative transfer from the L1, which is a common reason for error in the L2 literature ([21, 24]).

Regarding the production test result, it is clear that learners have acquired two different ranges of VOT velar aspirated and unaspirated stops. This is confirmed by inferential statistics applied to two data sets, which show a significant difference between the two mean values. However, we notice that our participants produced English velars [k] and [k❸] with 66.34 and 85.15 ms VOTs, respectively. The difference between the mean VOTs of native speakers of English and KA learners is not significantly different for [k^h] (p>.4). This means they have learnt voiceless aspirated stops. However, their VOTs in unaspirated stops are different from native English speakers' VOTs. But the important thing is that their VOTs for [k] are also different from their VOTs for [k^h]. Thus, they have learnt to differentiate between aspirated and unaspirated velar stops, but their unaspirated velar stop is not native-like. But in the opinion of [10], if a group of learners can realize the difference between two sounds, they have learnt this pair of sounds. Our conclusion about this sound is that KA learners have learnt unaspirated sound, but their sound is not native-like.

The result of /g/ is very scattered in nature. It cannot be analyzed or described in a single generalization. The reason is that the learners produced some tokens of /g/ with positive VOT and some with negative VOT. Therefore, we analyzed both in different ways. A one sample t-test was applied to compare the positive mean VOT of /g/ in the productions of KA English learners with the VOTs of /g/ by native speakers. The test confirmed that there was no significant difference between native speakers of English and Kuwaiti learners (p>.5). Those tokens which were produced with positive VOT are quite native-like. 198 out of 318 tokens were produced with a positive VOT. In other words, 62.24% of the tokens of /g/ velar stops were produced by KA learners, like that of English native speakers. Furthermore, 90 tokens were produced with a negative VOT of -74 ms pre-voicing. KA has velar stops with a negative VOT of -68.36 ms pre-voicing. Both values are closer to each other. These students transferred L1 VOT values into English. The remaining 30 out of 318 tokens were produced with zero VOT. This indicates learning because this stage is between the native-like category and L1 category. We can therefore conclude that in the production test results for /g/, we find 62% learning, almost 28% L1 transfer, and the remaining 10% tokens reflect that students are on the way to learning English velar stop /g/.

5.2. OT Based Analysis

5.2.1. OT Discussion

After studying the production test data of KA learners' English velar stops (keys, skis, geese), it was clearly seen that the learners had acquired two different ranges of VOT for velar stops as aspirated in "keys" and unaspirated in "skis." On the other hand, the KA learners also succeeded in distinguishing between aspirated and unaspirated velar stops. For this reason, the aspirated and unaspirated velar stops will not be tackled by OT formalism in this discussion because the KA learners have already adopted the constraint ranking of L2 English, which is in the same way as English native speakers.

The voiced velar stop /g/ will be only analyzed in this section by using classic OT constraint-based phonology paradigm. The finding of /g/ is divided into three groups of the KA learners: 1) a number of participants are native-like since they are producing /g/ with positive VOT like the English monolinguals, 2) another group of participants are producing /g/ with negative VOT like the KA monolinguals, and 3) a small number of participants in this group are producing /g/ with zero VOT, following neither English monolinguals nor KA monolinguals. Both the positive and negative VOT groups will be included in the OT analyses, whereas the zero VOT group will be excluded because they have not followed any grammar rules in their native and non-native languages.

English grammar and KA grammar each have their own constraint rankings, meaning that the ranking for KA monolinguals is different from the ranking for English monolinguals. The KA learners who have successfully learnt English switched from Kuwaiti grammar to English grammar, but those who could not learn English simply have a negative transfer coming from their KA L1 constraint rankings, affecting their English acquisition.

There are two types of VOT ranges: negative and positive. Stops with negative VOT are called "truly voiced stops". For example, KA has truly voiced stops. The VOT for KA /k/ in the word "keef" is 57 ms, which may be considered unaspirated, meaning that the KA velar stop is produced with a VOT of aspirated stops, but the quantity of aspiration is smaller than that observed in the English velar stop.

English /g/ has a positive short lag VOT, quite like the unaspirated [k] of British English. Thus, when a KA speaker hears a stop with a short-lag VOT, s/he perceives it as voiceless (because for him/her, only a pre-voiced stop, i.e., that with negative VOT, can be voiced [or +voice]), but the same stop with a short-lag stop is perceived as phonologically voiced (or +voice) by a British native speaker.

In other words, a British native speaker has a constraint to perceive stops with phonetically short-lag VOT as phonologically voiced (or +voice), but a Kuwaiti listener perceives the same stop with short-lag VOT as voiceless (or -voice) because for a Kuwaiti, only a stop with negative VOT can be voiced (or +voice). Based on this, we say PERCEIVE SL-VOT (+voice) (read it as "perceive a stop with a phonetically short lag VOT as phonologically voiced") is highly ranked in English, but in KA, PERCEIVE SL-VOT (-voice, i.e., "perceive a stop with a phonetically short lag VOT as phonologically voiceless") is highly ranked. Those learners of our study who perceive an English /g/ (with a positive short lag VOT) as voiced or (+voice), are following British English constraints, but those who perceive it (a stop with a short-lag VOT) as a voiceless or (-voice) stop are following KA constraint ranking. Those who are following Kuwaiti L1 ranking in the perception of British English /g/ are strongly under the influence of L1 grammar, but those who are following British English constraints in perceiving English /g/ are successful L2 learners because they have adopted British English ranking for an English stop.

5.2.2. OT Constraints and Table

In this section, the result of English /g/ acquisition by KA leaners will be treated within the framework of OT using classic OT [36]. The following cases cover both constraints, definitions, and rankings for L1 KA and L2 English, meaning that the findings will be divided into either a native-like group, which is producing /g/ with positive VOT, or a non-native-like group, which is producing /g/ with negative VOT.

(i). OT Constraints Definitions

The following constraints are used in the OT tables to show the performance of KA leaners of the English velar stop /g/:

The definition of constraint "PERCEIVE SL-VOT [-voice]":

Short-lag VOT is perceived as a voiceless stop. Assign a violation mark if short-lag VOT is not perceived as a voiceless stop.

The definition of constraint"PERCEIVE SL-VOT [+voice]": Short-lag VOT is perceived as a voiced stop. Assign a violation mark if short-lag VOT is not perceived as a voiced stop.

The definition of constraint "IDENT-PLACE":

The output segment and its input correspondent must have identical values in the [place] feature.

The definition of constraint "IDENT-VOICE":

The output segment and its input correspondent must have identical values in the [voice] feature.

(ii). Constraint Ranking for L1 KA

As a result of the following ranking, an English /g/ that is a short-lag stop with positive VOT will be perceived as a voiceless or -voice stop, which is [k].

PERCEIVE SL-VOT [-voice], IDENT- IO [PLACE], IDENT- IO [VOICE] >> PERCEIVE SL-VOT [+voice]

Table 9 shows a conflict between the highest-ranked constraint, PERCEIVE SL-VOT [-voice], and both the faithfulness constraints (IDENT-IO [PLACE], IDENT-IO [VOICE], and PERCEIVE SL-VOT [+voice]. This means that PERCEIVE SL-VOT [-voice] dominates PERCEIVE SL-VOT [+voice], supporting the fact that when KA leaners listen to the stop /g/ with a short-lag positive VOT, they perceive it as a voiceless /k/ as in "skis." Also, the highest-ranked constraint precedes the faithfulness constraints to favor the winning candidate.

The underlying form for the following table is /g/, and the following rankings are appropriate for this case:

PERCEIVE SL-VOT [-voice], IDENT- IO [PLACE], IDENT- IO [VOICE] >> PERCEIVE SL-VOT [+voice]

Table 9. */♑/→[k].*

/g/→[k]	PERCEIVE SL-VOT [-voice]	INDENT-IO [PLACE]	INDENT-IO [VOICE]	PERCEIVE SL-VOT [+voice]
a. /g/	*!			
☞ b. /k/				*
c. /d/	*!	*		
d. /t/		*		*

The candidate /g/ is ruled out due to the fatal violation of the highest-ranked constraint PERCEIVE SL-VOT [-voice], showing that this constraint demands a KA learner to perceive a stop with a phonetically short-lag VOT as voiceless. Candidate /k/ emerges as a winner because it satisfies the highly ranked constraints. Other candidates (/d/, and /t/) are defeated on account of the violation of the faithfulness constraints of the IDENT family, meaning that the candidates must not change place, or voice, respectively.

(iii). Constraint Ranking for L2 English

The following table 10 presents an argument for ranking PERCEIVE SL-VOT [+voice] above the faithfulness constraints (IDENT- IO [PLACE], IDENT- IO [VOICE]), and PERCEIVE SL-VOT [-voice]. This means that the constraint PERCEIVE SL-VOT [+voice] must be ranked higher than the constraint PERCEIVE SL-VOT [-voice], supporting the fact that KA learners perceive a stop with a phonetically short-lag VOT as a phonologically voiced /g/. Also, the highest-ranked constraint is more highly ranked than the faithfulness constraints to support the same fact for the winning candidate.

The underlying form for the following table is /g/, and the following rankings are appropriate for this case:

PERCEIVE SL-VOT [+voice], IDENT- IO [PLACE], IDENT- IO [VOICE] >> PERCEIVE SL-VOT [-voice]

Table 10. */g/→[g].*

/g/→[g]	PERCEIVE SL-VOT [+voice]	INDENT-IO [PLACE]	INDENT-IO [[VOICE]	PERCEIVE SL-VOT [-voice]
☞ a. /g/				*
b. /k/	*!		*	
c. /d/		*		*
d. /t/	*!	*	*	

The candidate /g/ is chosen as the winner because it satisfies the highest-ranked constraint PERCEIVE SL-VOT [+voice], supporting the fact that a KA learner perceives a stop with a phonetically short-lag VOT as voiceless. Candidate /k/ emerges as a losing candidate because it disfavors the highly ranked constraints. Other candidates (/d/, and /t/) are violated by the faithfulness constraints of the IDENT family, meaning that the candidates must not change their place or voice.

6. Conclusion

This paper was based on the perception and production of English velar stops. A large group of Kuwaiti learners of English perceived and produced velar stops. Their discrimination between voiced and voiceless velar stops was excellent, but their identification of [g] was weak. Their identification of the voiceless velar stop /k/ was also native-like. In production, they were native-like in the aspirated [k^h]. In the production of the unaspirated voiceless velar stop [k], they were not native-like, but they had developed an understanding of this allophone of English. With [g], some students were native-like, some were still learning, and some only relied on L1 transfer. For this reason, classic OT was used to treat the [g] and provide phonological evidence, showing that some of the KA participants have learnt English /g/ and have adopted English L2 constraints, whereas some have not learnt this and still use KA L1 constraints for English.

7. Recommendations

The current study has focused on the acquisition of allophonic contrast in velar stops by adult KA learners of English.

Firstly, this study can be extended to other English plosives like coronal /t/ and labial /p/ which also have similar aspiration contrast in English. Such a study could be more interesting in the perspective of the fact that English coronal /t/ is alveolar and KA coronal /t/ is dental. Similarly, English /p/ is a classical example of L2 learning difficulty for Arab learners. Normally, previous studies have focused on the acquisition of /p/ phoneme by Arab learners but not on the study of allophonic variance.

Secondly, the same study may be extended to other phonemes of English that have allophonic variance. The English lateral is another example of allophonic variance. English lateral is produced as dark in the coda position, but it is produced as a clear lateral in other positions. A study of the acquisition of allophonic variance in English lateral by KA learners may have a very solid contribution towards the literature on the acquisition of allophonic variance in L2.

Thirdly, the same study may be extended to other adult learners of English. The findings of such a study will further substantiate the previous findings in this field of study.

References

[1] Alanazi, S. M. (2018). The acquisition of VOT of [b d g] in the context of Saudi learners of English. مجلة التربية *(الأزهر): والاجتماعية والنفسية التربوية للبحوث محكمة علمية* [a refereed scientific journal for educational, psychological and social research], *37* (178), 659-678.

[2] Adra, M. A. (1999). *Identity effects and opacity in Syrian Arabic: An optimality theory analysis*. PhD thesis, University of Illinois at Urbana Champaign.

[3] Aldaihani, S. M. (2014). *Phonological processes in Kuwaiti Arabic*. PhD dissertation University of Essex, United Kingdom.

[4] Al-Mohanna, F. (1998). *Syllabification and metrification in urban Hijazi Arabic: Between rules and constraints*. PhD thesis, University of Essex, UK.

[5] Al-Omar, M. (2011). *A comparative study of phonological processes in Syrian and Jordanian Arabic: An optimality theoretic approach*, PhD thesis, University of Essex, UK.

[6] Bermúdez-Otero, R. (1999). *Constraint interaction in language change: Quantity in English and Germanic*. PhD dissertation, University of Manchester.

[7] Boersma, P., & Weenink, D. (2020). Praat: Doing phonetics by computer (Version 6.0. 52).

[8] Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris publications. https://doi.org/10.1515/9783110872637.fm

[9] Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, London: Harper & Row.

[10] 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. New York: Timonium, MD.

[11] Harris, J. (1994). *English sound structure*. Oxford: Blackwell.

[12] Honeybone, P. (2005). *Diachronic evidence in segmental phonology: The case of laryngeal specifications*. Berlin: Mouton de Gruyter.

[13] Kabrah, R. S. (2004). *Opacity and transparency in the phonology of Makkan Arabic: A stratal optimality-theoretic approach,* PhD thesis, Boston University. USA.

[14] Kager, R. (1999). *Optimality theory*. New York: Cambridge University Press.

[15] Kehoe, M. M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German–Spanish children. *Bilingualism: Language and Cognition*, *7* (1), 71-88.

[16] Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*: Singular San Diego, CA.

[17] Kulikov, V. (2016, March). Voicing in Qatari Arabic: Evidence for pre-voicing and aspiration. In *Qatar Foundation Annual Research Conference Proceedings Volume 2016 Issue 1* (Vol. 2016, No. 1, p. SSHAPP2330). Hamad bin Khalifa University Press (HBKU Press).

[18] Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D.

[19] Kiparsky, P. (2000). Opacity and cyclicity, *Linguistic review*, *17* (2/4), 351-366.

[20] Kisseberth, C. (1970). On the functional unity of phonological rules, *Linguistic Inquiry*, 1, pp. 291-306.

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

[22] Larson-Hall, J. (2016). A guide to doing statistics in second language research using SPSS. London: Routledge.

[23] Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20* (3), 384-422.

[24] Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

[25] McCarthy, J. (1997). Process-specific constraints in optimality theory. *Linguistic Inquiry* 28, 231-251.

[26] McCarthy, J. (2000). Harmonic serialism and parallelism. In Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-yung Kim, eds., *Proceedings of the North East Linguistics Society* 30. Amherst, MA: GLSA. pp. 501-524.

[27] McCarthy, J. (2002). On targeted constraints and cluster simplification. *Phonology* 19, 273-292.

[28] McCarthy, J. (2010c). Studying GEN. *Journal of the Phonetic Society of Japan* 13 (2), 3-12.

[29] McCarthy, J., & Prince, A. (1990). Prosodic morphology and templatic morphology, in M. Eid, & McCarthy, J. (eds.) *Perspectives on Arabic Linguistics*, pp. 1-54, Amsterdam: John Benjamins.

[30] McCarthy, J., & Prince, A. (1993a). Generalized alignment, in Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology*, pp. 79–153. Dordrecht: Kluwer.

[31] McCarthy, J., & Prince, A. (1993b). *Prosodic Morphology I: Constraint interaction and satisfaction,* Technical Report #3, Rutgers University Center for Cognitive Science.

[32] McCarthy, J., & Prince, A. (1995). Faithfulness and reduplicative identity, in J. Beckman, L. W. Dickney, & S. Urbanczyk. (eds.), *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics 18*. Amherst, MA: Graduate Linguistic Student Association.

[33] Mustafawi, E. (2006). *An optimality theoretic approach to variable consonantal alternations in Qatari Arabic*, PhD thesis. The University of Ottawa.

[34] Nasukawa, K. (2010). Place-dependent VOT in L2 acquisition. In Matthew T. Prior et al. (Eds.), *Selected Proceedings of the 2008 Second Language Research Forum,* pp. 198-210. Somerville: Cascadilla Proceedings Project.

[35] Nelson & P. Foulkes (eds.), *Leeds Working Papers in Linguistics and Phonetics*, 8, 95– 122.

[36] Prince, A. and Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA, & Oxford: Blackwell. [Revision of 1993 technical report, Rutgers University Centre for Cognitive Science. Available on Rutgers Optimality Archive, ROA-537].

[37] Rakhieh, B. A. (2009). *The phonology of Ma'ani Arabic: Stratal or parallel OT*. PhD thesis, University of Essex.

[38] Roca, I. and Johnson, W. (2007). *A course in phonology*. Oxford: Blackwell.

[39] Rose, S. (2000). Doubled verbs and syncope resistance in Iraqi Arabic: Not antigemination, In *14th Arabic Linguistic Symposium, UC Berkeley*.

[40] Scholfield, P. (1995). *Quantifying language: A researcher's and teacher's guide to gathering language data and reducing it to figures*. Clevendon/Philadelphia: Multilingual Matters Ltd.

[41] Shimizu, K. (2010). Acoustic analysis of Korean L2 learners in the acquisition of English and Japanese stop voicing contrasts. In *A paper presented in Thirteenth Australasian International Conference on Speech Science and Technology*.

[42] Simon, E. (2009). Acquiring a new second language contrast: an analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research*, *25* (3), 377-408.

[43] Syed, N. A. (2014). Why Pakistani learners cannot acquire English voiced stops? *Journal of Research in Social Sciences; Islamabad*, *1* (1), 19-30.