VOICE ONSET TIME (VOT) IN KURDISH EFL LEARNERS

Lureen Ibrahim Naser

Department of English Language, College of Basic Education, University of Duhok, Kurdistan Region, Iraq.

lureen.nasar@uod.ac

ABSTRACT

This paper examines the acquisition of English stops by Bahdini Kurdish learners. Kurdish utilizes a three-way category of voiced, voiceless aspirated and unaspirated stops. Kurdish voiced stops are pre-voiced, while those of English have a short-lag VOT, characteristic of Kurdish voiceless unaspirated stops. Kurdish also has voiceless aspirated stops with shorter VOT than those of English. These differences require learners to restructure their existing categories which is more challenging than creating new categories [1, 2].

The sample consisted of two monolingual control groups of (40) English and Kurdish speakers and two study groups of (60) beginner and advanced Kurdish EFL learners. Two lists of monosyllabic words are constructed, in English and Kurdish, with initial prevocalic stops providing 4140 tokens for analysis. Results show that learners lengthened their VOT towards English as they progress from beginner to advanced stage. Their VOT patterns, however, remain intermediate between both languages.

KEYWORDS: Kurdish EFL Learners, VOT, English stops, Kurdish stops.

1. INTRODUCTION

There are several acoustic signals in speech that help us to detect a foreign accent. Voice Onset Time (VOT) in stops has long been recognized as a significant feature in distinguishing accented speech and an important parameter susceptible to measurement and analysis.

In second language (L2) acquisition, learners' productions in the target language (TL) are constrained by the perceptual experience of their first language (L1), according to the Perceptual Assimilation Model-L2 (PAM-L2) [3]. The perception of learners is adjusted constantly for both L1 and L2. The Speech Learning Model (SLM) [2] postulates that learners create phonetic categories which include long-term memory representations created through exposure to the acoustic-phonetic cues of the language. The SLM hypothesizes that, based on the input learners receive, they start developing sub systems for the language they are trying to learn to encompass the new realization rules of the acquired language. Learners use this subsystem when they are in the TL processing mode. However, learners commonly deviate from the phonetic norms of the TL. The ability to form new phonetic categories decreases due to the fully developed and stabilized phonetic system of the L1, especially with limited L2 input [1]. When both L1 and TL share similar but not identical phonetic categories, it is harder for learners to establish or even adjust or expand those of their L1 to accommodate these slight changes [2]. Thus, small differences may be more problematic for learners who need to re-structure previously established phonetic categories to accommodate the features of the new language.

An example that may support this hypothesis would be the case of this study concerning the acquisition of English stops by Kurdish learners. Both languages have voiced stops, however, there are different acoustic properties attributed to the voiced stops of each language with reference to VOT. English voiced stops are produced with short lag VOT, characteristic of Bahdini Kurdish voiceless unaspirated stops, while Kurdish voiced stops are produced with voicing lead, similar to some other Indo-Iranian languages [4]. These different phonetic details are expected to be challenging for Kurdish learners and may make it harder for them to achieve native-like VOT. Unlike English, Kurdish contrasts voiceless aspirated and unaspirated stops. So, the voiceless stops of both languages are also recorded to check if they occupy different areas along the VOT continuum. Accordingly, the aim of this paper is to answer the following questions:

- Do Kurdish EFL learners produce English voiced stops with a negative VOT characteristic of Kurdish or a short lag VOT characteristic of English?
- Can they adjust their two categories of voiceless stops (aspirated and unaspirated) to that of one English phonological category of voiceless stops?
- Does the proficiency level of learners affect their performance?
- Are the variables of laryngeal state, place of articulation, vowel height or word duration significant?

2. METHODOLOGY

2.1. Stimuli

Tow lists of monosyllabic words were constructed in English and Kurdish. The phonological voiced category of stops was represented by initial /b/, /d/ and /g/ and the phonological voiceless category of stops was represented by initial aspirated and unaspirated (only for Kurdish) /p/, /t/ and /k/. The lists included two example words for each stop, one followed by a high vowel and one by a low vowel.

2.2. Participants

Four groups of participants were chosen for this study. Two control groups of (30) monolingual adult native speakers of Bahdini Kurdish and (10) monolingual adult native speakers of English. They were aged between 20-50 years. The experimental groups were (60) adult Kurdish EFL learners from two proficiency levels. Both groups were university students. The participants had all studied English but differed in the amount of academic input they received in their study starting from one year to four years. They were aged between 20-30 years of age. All groups had equal numbers of males and females.

2.3. Acoustic procedures

Monolingual English and Kurdish participants were asked to read a list of words with initial voiced and voiceless stops, constructed by the researcher, each in their native language. The experimental groups of Kurdish EFL learners were asked to read the same list read by English native speakers. Each word was repeated (3) times, which provided 3 tokens for each word. Participants were asked to repeat a word when pronounced incorrectly. Productions were recorded using a portable Zoom H1n digital audio recorder with a built-in microphone. All the recordings were in uncompressed WAV format, sampled at 44.1 kHz with 16 bits per sample. They were later coded and segmented in Audacity into separate WAV files, one file for each word in the list to be ready for analysis by *Dr.VOT* software [5], which is a package for automatic measurement of VOT. All the resulting acoustic measurements of VOT were visually reviewed and manually corrected in *Praat* [6].

2.4. Statistical procedures

Statistical analyses were carried out using R [7]. To investigate the effect of a group of factors on VOT, the LME model, using the *lme4* package [8] and the *lmerTest* package [9], was adopted. Each studied group had a different LME model. Fixed Factors in the model included Place of Articulation (POA) (bilabials, dental, velar), Laryngeal State (voiceless aspirated, voiceless unaspirated and voiced), Vowel Height and carrier word Duration and a by-Speaker random factor with an alpha level $p = .05$. A post-hoc Tukey's test [10] was carried out to show the direction of difference between the three different levels POA. Another post-hoc Dunnett's test, using *DescTools* package [11], was used to test statistical differences between all studied groups and to conclude whether VOT values produced by any of the research groups, Kurdish Learners G1 and G2, are closer to Kurdish or English.

3. RESULTS

Table 1 shows the mean (M), standard deviation (Std.D) and range (R) of VOT values of word initial stops for all studied groups. Voiced stops are characterized by lead for all groups except for English /d/ and /g/, which have a short lag. Kurdish voicing lead, however, tends to be shorter in Kurdish than the other groups. Kurdish voiceless unaspirated stops are produced with a short lag. Voiceless aspirated stops, on the other hand, are all produced with a long lag. Voiceless stops show a pattern across all groups in which bilabials have the shortest VOT, followed by dentals then velars, which is the universal pattern [12]. This is also evident in the results of the post-hoc Tukey test.

Kurdish						
	/b/	**/d/**	**/g/**	**/kʰ/**	**/pʰ/**	**/tʰ/**
Mean	-114	-119	-114	86	57	62
Std.D	32	34	33	21	15	20
Range	-198:-51	-214:-52	-209:-42	50:155	36:103	31:121
				/k/	**/p/**	**/t/**
Mean				25	11	16
Std.D				5	3	6
Range				14:39	5:21	5:31
English						
	/b/	**/d/**	**/g/**	**/kʰ/**	**/pʰ/**	**/tʰ/**
Mean	-12	16	21	120	90	99
Std.D	36	13	17	24	21	27
Range	-99:12	-54:24	-56:38	50:159	50:120	31:145
KEFL1						
	/b/	**/d/**	**/g/**	**/kʰ/**	**/pʰ/**	**/tʰ/**
Mean	-108	-109	-103	91	62	70
Std.D	36	38	41	26	24	26
Range	-197:11	-203:19	-196:23	35:157	23:118	20:136
KEFL2						
	/b/	**/d/**	**/g/**	**/kʰ/**	**/pʰ/**	**/tʰ/**
Mean	-82	-72	-63	88	55	68
Std.D	35	46	48	21	19	20
Range	-170:11	-178:23	-172:38	33:137	20:108	27:126

Table 1: Range (R), Mean (M) and Standard Deviation (Std. D) of VOT values of the stop categories for all studied groups measured in milliseconds (ms).

To conclude whether VOT values produced by any of the research groups, Kurdish Learners G1 and G2, are closer to Kurdish or English, a post-hoc Dunnett's test was used. It showed significant differences between all four groups. The results of the test concluded that Kurdish Learners G1's (KEFL1) mean VOTs are closer to Kurdish (-3.73, p = .574) than to English (-71.8, p < .001). The Kurdish Learners G2's (KEFL2) mean VOTs, however, reveal change towards English (-57.9, p < .001) and away from Kurdish (-10.15, p = .006). Based on this, we can conclude that Kurdish learners have developed towards English, but still produce English voiced stops with notably long negative VOT.

3.1. VOT and laryngeal state

Laryngeal state of initial stops showed a significant effect on their VOT in all studied groups (F (2, 1568) = 51.34; p < .001) for Kurdish, (F (1, 267) = 7.86; p = .005) for English, (F (1, 998) = 160.99; p < .001) for Kurdish learners G1 and (F (1, 1014) = 145.09; p < .001) for Kurdish learners G2. The two laryngeal categories of voiced and voiceless states had significantly different VOT values, (p < .001) for all studied groups with voiceless stops having longer VOT values than voiced ones, as evident in Figure 1.

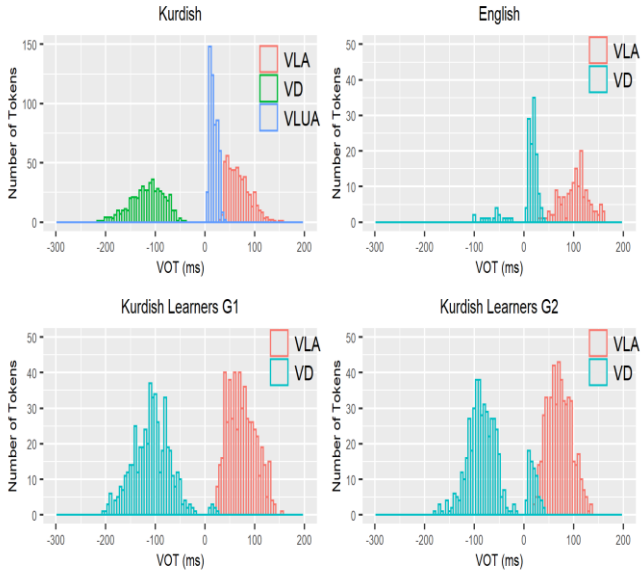

Figure 1: Histogram of the distribution of VOT values of stop categories for all studied groups.

3.2. Place of articulation

Place of articulation of the initial stops showed a significant effect on VOT only for the English and Kurdish G2 groups (F (2, 1557) = 5.26; p = .005) for Kurdish, (F (2, 267) = 1.52; p = 221) for English, (F (2, 996) = 12.59; p < .001) for Kurdish learners G1 and (F (2, 1007) = 1.19; p = .304) for Kurdish learners G2.

The distinction in VOT was evident between the three places of articulation. Bilabials and Velars differed from each other significantly in all of the studied groups (p < .001), except for the English group (p = .072) and Kurdish learner G2 group (p = .262). However, bilabials didn't show such high significant difference in VOT with dentals except for the Kurdish (p = .003) and Kurdish learners G1 group (p = .002). English and Kurdish learners G2 and Kurdish had the following values respectively; (p = .377) and (p = .025). These results are visualized in Figure 2.

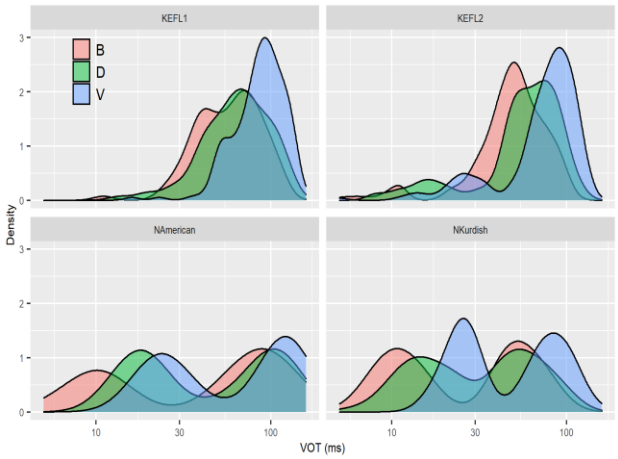

Figure 2: Density plots of the distribution of VOT values based on POA for all groups.

To assess the direction of difference between the three places of articulation, a post-hoc Tukey test [11] was performed. Mean VOT values of voiceless stops show a pattern across all groups in which bilabials have the shortest VOT, followed by dentals then velars.

3.3. Speech rate (Duration)

Duration of word carrier had a statistically significant effect on the VOT of Kurdish, Kurdish learners G1 and Kurdish learners G2 groups (F (1, 468) = 53.71; p <.001) for Kurdish, (F (1, 863) = 38.80; p =<.001) for G1 and (F (1, 1027) = 33.06; p =<.001) for G2. The VOT of the English group, however, didn't show a statistically significant effect of word duration (F (1, 273) = 0.30; p =.582).

3.4. Post vowel height

Height of the following vowel only showed a significant effect on VOT of initial stops for native Kurdish but not for the other studied groups with: F (1, 1558) = 8.64; p =.003 for Kurdish, F (1, 266) = 2.25; p = .134 for English, F (1, 994) = 0.54; p = .464 for Kurdish learners G1 and F (1, 1010) = 0.03; p = .863 for Kurdish learners G2.

As displayed in Figure 3, the distinction in VOT between high and low vowels was only statistically significant for the Kurdish and learners G1 groups with a $p < .001$, $p = .337$, $p = .011$, $p = .564$ for Kurdish, English, learners G1 and G2 respectively.

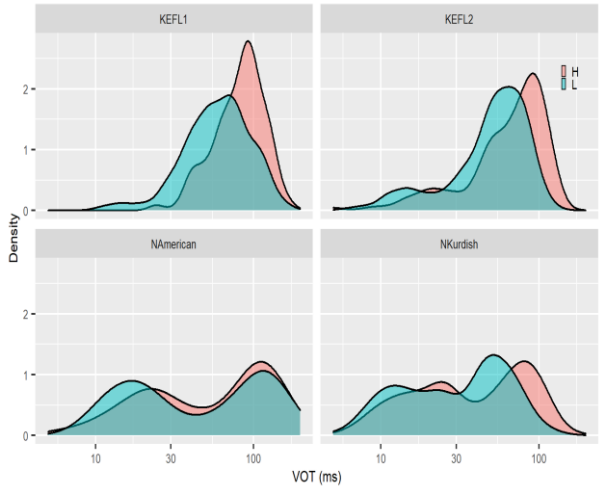

Figure 3: Density plots of the distribution of VOT based on vocalic context for all groups.

4. SUMMARY AND DISCUSSION

Extensive research has been carried out to investigate the role of VOT in L2 stop productions. Learners of varied native language backgrounds exhibit different measures of TL VOT. This paper examines how L1 Kurdish EFL learners produce VOT for the voicing contrast between English voiced and voiceless stops. Kurdish contrasts pre-voiced voiced stops, short-lag unaspirated stops, and long-lag aspirated stops, unlike English which contrasts long-lag voiceless stops with short-lag voiced stops. The three-way voicing contrast in L1 Kurdish is mapped onto a two-way contrast in L2 English, in which all three VOT categories are employed at the level of phonetic realization.

Results show that learners produce pre-voicing for English voiced stops, though less than in Kurdish. Long-lag VOT for voiceless aspirated stops in Kurdish and in EFL learners tends to be shorter than that of voiceless aspirated stops of English by native speakers. Kurdish EFL learners transfer their voiced stop lead VOT values to the TL and tend to produce English voiced stops with a negative VOT, especially in lower proficiency levels. This transfer, however, tends to be less for more advanced levels, whose VOT productions of voiced stops involve both negative and positive values. VOT values of the L1 unaspirated voiceless stop category (short lag) overlap with the L2 voiced stop category. Kurdish speakers have to learn to produce the voiced phoneme with an optional short-lag VOT variant.

Learning an L2 phoneme which is only phonetically different from one in L1, would only require learners to adjust their existing phonetic categories to embrace the subtle changes of the new L2 variant. In the case of this study, although Kurdish EFL learners show progress towards the target language, they still produce English voiced stops with notably long negative VOT values. They also produce English voiceless aspirated stops with less aspiration than the English group. This supports Flege's [2] view that small differences can be more challenging for learners. The inability to produce TL stops with accurate VOT leads to retaining an amount of foreign accent in their speech. It may be due to the following reasons, which represent general common concepts among L2 speech theories.

- L2 learning took place in a foreign language acquisition environment (FLA) where the L1 is the predominant language and is not used outside the classroom. This supports the notion that the ability to form new phonetic categories decreases with limited L2 input.
- Structure of L1 of learners has an effect on learning that of an L2, although phonetic or phonological similarities do not guarantee ease of learning [3]. According to the SLM [2], these learners had already created phonetic categories for their native Kurdish voiced stops with negative VOT values. When they try to learn English, they will only need to adjust that of their native language to include the subtle phonetic details of English voiced stops with a short lag. This is challenging, since the phonetic details of both languages exist in the same phonological space; meaning that voicing lead and short lag are both associated with voiced stops.
- The level of learners (beginner or advanced) determines the degree of effect of certain variables on their L2 production. Advanced learners, for instance, are expected to be able to distinguish phonetically similar structures between L1 and L2 more easily than beginners. Such similarity may create confusion in novice learners. In this study, advanced learners have shown some progress by producing English voiced stops with shorter negative VOT values, compared to the long negative VOT values produced by beginners.

5. ACKNOWLEDGMENTS

The author would like to thank Professor Jane Stuart-Smith at Glasgow University Laboratory of Phonetics (GULP) for her valuable feedback on earlier versions of the paper. Thanks to Dr. Clara Cohen for her statistical advice on R.

6. REFERENCES

1] Flege, J. E. 2003. Assessing constraints on second-language segmental production and perception. Phonetics and phonology in language comprehension and production: Differences and similarities, 6, 319-355.

[2] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. Speech perception and linguistic experience: Issues in cross-language research, 92, 233-277.

[3] Best, C. T., & Tyler, M. D. 2007. Nonnative and second-language speech. Language experience in second language speech learning: In honor of James Emil Flege, 17, 13.

[4] Hussain, Q. (2018). A typological study of voice onset time (VOT) in Indo-Iranian languages. *Journal of Phonetics* 71, 284–305.

[5] Shrem, Y., Goldrick, M., & Keshet, J. 2019. Dr. VOT: Measuring positive and negative voice onset time in the wild. Proc. Interspeech. arXiv preprint arXiv:1910.13255. https://arxiv.org/pdf/1910.13255.pdf

[6] Boersma, P. and Weenink, D. 2018. "Praat: Doing phonetics by computer [computer program]," Version 6.0.39. URL http://www.praat.org/.

[7] R Core Team 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[8] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

[9] Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. 2017. LmerTest package: Tests in linear mixed effects models. Journal of Statistical Software, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13 (R lmerTest package: version 2.0-36).

[10] Keselman, H. J., & Rogan, J. C. 1977. The Tukey multiple comparison test: 1953–1976. Psychological Bulletin, (845), 1050.

[11] Andri et mult. al. S 2022. *DescTools: Tools for Descriptive Statistics*. R package version 0.99.45, https://cran.r-project.org/package=DescTools.

[12] Cho, T. & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27, 207–229.