Lexical Representation of Noncanonical Forms: Evidence from Persian

Koorosh Ariyaee (koorosh.ariyaee@mail.utoronto.ca)
Department of Linguistics, University of Toronto
100 St. George St., Toronto, ON, M5S 3G3, Canada

Philip J. Monahan (philip.monahan@utoronto.ca)
Department of Language Studies, University of Toronto Scarborough
1265 Military Trail, Toronto ON, M1C 1A4, Canada

Jessamyn Schertz (jessamyn.schertz@utoronto.ca)
Department of Language Studies, University of Toronto Mississauga
3359 Mississauga Road North; Mississauga, ON L5L1C6, Canada

Abstract

The lexical representation of words with multiple pronunciation variants has been widely debated: while single storage accounts propose that all variants of a word are represented by a single, canonical, representation, multiple storage accounts include representations for different pronunciation variants. Previous work has provided evidence for the representation of noncanonical variants, consistent with multiple storage models; however, this work has focused on highly frequent noncanonical variants, which some single storage models propose can also be lexically represented. To test predictions of multiple storage models more rigorously, we examined the representation of low-frequency noncanonical variants of the Persian uvular stop [ɢ]. Results from three experiments support the multiple storage model, in that they provided evidence for lexical storage of low-frequency variants. Implications of these findings are discussed.

Keywords: noncanonical forms; phonetic processing; lexical representation; mental lexicon; speech perception; Persian

Introduction

Canonical pronunciations are the citation forms of words with no segmental changes. Such forms are not always observed in natural speech (see Ernestus, 2014); instead, noncanonical or reduced word pronunciations frequently appear in daily conversation (Johnson, 2004). Noncanonical forms can arise via gradient phonetic changes (e.g., natural variation in voice onset time and segment duration; Sumner et al., 2014), as well as when canonical forms undergo various processes, such as deletion ([kæmərɪə] → [kæmɹə] 'camera') or lenition ([bɛtəɹ] → [bɛɾəɹ] 'better'). The current research focuses on the latter, when noncanonical variants differ categorically from their canonical counterparts at the segmental level.

An ongoing debate relates to the representation of phonetic variants of the same word (e.g., Dell, 1986; Goldinger, 1998; Spinelli & Gros-Balthazard, 2007; Sumner & Samuel, 2009; Sumner, 2013; Soo & Babel, 2024). The central question is whether multiple pronunciation variants of the word, particularly noncanonical forms, are lexically represented. To this end, different models of speech representation have been developed, roughly classified into two categories: single and multiple storage models.

Single storage models differ from multiple storage accounts in two key assumptions. First, influenced by generative grammar (Chomsky & Halle, 1968), proponents of single storage models suggest that among multiple pronunciation forms of a given word, only the canonical variant is stored. That is, comprehension involves mapping various phonetic forms of the same word onto a single lexical representation (Caramazza, 1997; Connine et al., 1993; Dell, 1990; McClelland & Elman, 1986; Norris, 1994; Samuel & Larraza, 2015; Stevens, 2002). Within this framework, the canonical word variant is represented in the lexicon with all its features (Connine et al., 1993). A noncanonical variant as the input, which differs from the canonical variant by at least one feature, less effectively activates the target word, compared to the canonical input (Blumstein, 2004). As such, the processing time and activation of the intended lexical representation varies depending on the degree of mismatch between the input and the representation (Marslen-Wilson et al., 1996). As the mismatch between the input and represented form increases, the activation level of the represented form decreases: the less similar an input is to the represented form, the slower the reaction time (Bölte & Coenen, 2002). Single storage models predict that the canonical variant of a word is processed faster than its noncanonical variants, because features of the canonical input match those of the intended representation. This alignment of features between the input and the representation leads to faster recognition.

Another assumption of single storage models, setting them apart from multiple storage accounts, is that they do not consider the role of occurrence (relative) frequency of a pronunciation variant in its recognition. Relative or variant frequency is the occurrence frequency of a phonetic variant of a word (e.g., [kæmərɪə] 'camera') relative to the occurrence frequency of other phonetic variants of the same word (e.g., [kæmɹə]). According to single storage models, differences in reaction time (RT) are based on the canonicity of those variants (Dufour et al., 2022; Levelt et al., 1999; Luce et al., 2000). For instance, in the case of English schwa deletion, the noncanonical variant of *camera* (i.e., [kæmɹə]) should be recognized more slowly than an input with a retained schwa (i.e., [kæmərɪə]). This is because the unreduced variant matches the represented form more than the reduced input, regardless of the relative frequencies of [kæmərɪə] and [kæmɹə]. As such,

the degree of mismatch between noncanonical forms and their canonical counterpart determines their processing time, and not the occurrence frequency of variants. This view implies that canonical word forms have a privileged lexical status; they are the only represented form, and thus noncanonical variants are processed based on phonetic similarities to their canonical counterparts, and not based on their own properties, such as occurrence frequencies (see Ranbom & Connine, 2007).

Unlike single storage systems, multiple storage models posit that different pronunciation variants are lexically stored per word. That is, besides the canonical form, their categorically distinct noncanonical form(s) are also stored in the lexicon (e.g., McQueen et al., 2003; Pitt, 2009; Sebastián-Gallés et al., 2009). Such models assume that there is one representation separately stored per categorically distinct phonetic variant (Bürki et al., 2010). For instance, each of the two distinct pronunciation variants of the word *camera* (i.e., [kæmərɪə], [kæmɪə]) is represented in the mental lexicon. Proponents of multiple storage systems posit that the frequency of each variant affects the processing time of that variant. Based on results from various experiments, proponents of multiple storage models (e.g., Bürki & Gaskell, 2012) posit that higher frequency variants elicit faster RTs, suggesting an inverse relationship between a variant's frequency and its RT. For instance, Ranbom and Connine (2007) examined the representation of variants of English words with two pronunciation forms, such as [kaʊ**nt**i] and [kaʊ**ɾ**i] as canonical and noncanonical forms of the word *county* in American English. They conducted a rating experiment to calculate the variants' relative frequencies, as well as lexical decision and priming experiments to measure their RTs. High-frequency variants resulted in faster RTs than low-frequency variants.

Similarly, Bürki et al., (2010) tested French schwa words with two pronunciation variants (e.g., canonical: [ʁəkɛ̃], noncanonical: [ʁkɛ̃] *requin* 'shark'). The relationship of variant frequency and RT, as collected from rating and priming experiments, revealed that higher-frequency variants elicited faster RTs than lower-frequency variants. Thus, variant frequency affects processing time and as such, provides insight into their representation. They also suggested that each variant frequency is stored, and variants differ from one other in their frequency; hence, the frequency of each variant distinguishes itself from other variants during speech processing (Bürki et al., 2010).

Using rating and lexical decision experiments, Brand and Ernestus (2018) examined the representation of noncanonical variants and their frequencies. They tested whether listeners' own level of exposure to each variant influenced processing phonetic forms of words. To do so, three groups of participants with three levels of French proficiency (i.e., native, advanced learner, beginner learner) completed a rating and lexical decision experiment involving French schwa words with two pronunciation forms (e.g., canonical: [ʒənu], noncanonical: [ʒnu], *genou* 'knee'). For native speakers, RTs and accuracy of canonical and noncanonical variants strongly correlated with their relative frequencies. For advanced learners, accuracy scores and RTs correlated with their relative frequencies; however, in the beginner French learner group, RTs and accuracies did not correlate with their relative frequencies for noncanonical variants. Brand and Ernestus concluded that these results indicate that different levels of experience with variants affect RTs and recognition of these forms between groups (see also Sumner & Samuel, 2009). Crucially, they suggested that native French speakers and advanced learners had represented the noncanonical variants and the frequencies of said variants—as evidenced by the correlation between RT and relative frequencies; however, since there was no such correlation between RT and noncanonical variant frequency in beginner learners, there was no evidence that they had stored these variants and their frequencies. Comparing the beginner learners to the other groups, they concluded that their study provided evidence for the representation of relative frequencies of phonetic variants and consequently, the variants themselves in the advanced learner and native speaker groups.

Taken together, these studies suggest that if a given variant is stored, there should be an inverse relationship between relative frequency and RTs of that variant. Obtaining such a relationship for a variant can indicate that the occurrence frequency of variants plays a crucial role in their recognition and representation—and not necessarily their canonicity.

Goals and Motivations

Previous studies providing evidence for multiple storage models use data where noncanonical variants are more frequent than their canonical counterparts; for example, Ranbom and Connine (2007) test the representation of noncanonical variants with nasal flaps in American English (e.g., [kaʊ**ɾ**i] 'county') that cover around 80% of all productions with the /nt/ sequence (e.g., [kaʊ**nt**i] as the canonical form). Similarly, Bürki et al., (2010) and Brand and Ernestus (2018), examine the representation of noncanonical forms with deleted schwas in French, and Bürki and Gaskell (2012) investigate similar patterns in English; in both languages, the forms with deleted schwas (i.e., noncanonical) are more common than their canonical counterparts (see Fougeron et al., 2001). Examining patterns of variation in which noncanonical variants are high-frequency can pose challenges for predictions of multiple storage models. Some proponents of single storage accounts (e.g., Levelt, 1989) posit that while the mental lexicon stores only the canonical variant, it can also exceptionally store noncanonical forms that are considerably more frequent than their canonical counterparts. Thus, evidence for the storage of high-frequency noncanonical forms, while in line with multiple storage accounts, is also consistent with predictions of some single storage models. Hence, testing only high-frequency noncanonical variants may not differentiate predictions of single storage models from multiple storage accounts, as both yield the same predictions for the representation of high-frequency noncanonical forms.

A more ideal test to distinguish predictions of these two models, that is, specifically test predictions of multiple storage models to the exclusion of single storage models, is a

situation where noncanonical variants are less frequent than their canonical counterparts. Evidence for the storage of low-frequency noncanonical variants would not be consistent with predictions of single storage accounts, as these models assume low-frequency noncanonical variants are not stored. Thus, evidence for the representation of low-frequency noncanonical variants would support predictions of multiple storage models only. Such an outcome would indicate that, unlike predictions of single storage systems, a noncanonical variant can be lexically represented and that variant does not necessarily need to be high-frequency to be lexically stored.

The goal of this study is to provide a stronger test of multiple storage models than has been carried out in previous work by examining the storage status of low-frequency noncanonical variants. To address this question, the current study examines whether multiple phonetic forms of words with the uvular phoneme in Persian are stored. Previous work has shown that words with a uvular stop in their canonical form (e.g. [moɢol] 'Mongol') can also be pronounced with a uvular approximant ([moʁ̞ol]), but this noncanonical form is less frequent (Ariyaee, 2021). This pattern offers an excellent test for predictions of multiple storage accounts. Another contribution is that it extends research on the status of variation in lexical storage to a lesser-studied language. Previous work on the topic has focused on a limited set of languages and phenomena (Bürki, 2018). Increasing the breadth of tested languages and patterns is critical to ensure the broader generalizability of findings relevant to lexical storage models.

To explore the representation of these noncanonical variants in Persian, via the Gorilla platform (Anwyl-Irvine et al., 2020), we conducted three online experiments: a production, a rating, and a lexical decision experiment. First, a production study and an acceptability rating study were used to establish the relative frequency of the variants in different word positions (i.e., initial, medial, final). Then, a lexical decision experiment tested whether RTs were influenced by the relative frequency of variants in different positions, as predicted by a multiple storage model.

Production experiment

To address the research question, a preliminary step was to examine the occurrence rates of variants of the uvular phoneme in different positions. One method to obtain such information is using corpora of spontaneous speech (Bell et al., 2009; Purse et al., 2022; Tamminga, 2019); however, the current corpora of Spoken Persian (e.g., Mohammadi, 2019) use Persian orthographic symbols and do not provide phonetic transcriptions. The allophonic variants of the Persian uvular phoneme are not represented in Persian orthography. Therefore, these corpora do not capture the sub-orthographic, phonetic variation of the uvular phoneme in Spoken Persian. These corpora are also text-based and do not contain audio files to be used for measuring variant occurrence rates. Given these limitations, a production experiment was conducted to gauge the phonetic variation of the uvular consonant, establish the relative frequencies of its variants and examine whether noncanonical variants have lower occurrence rates than their canonical counterpart.

Methods

Sixty-six Persian speakers (35 f, 31 m; mean age = 33 years) completed the study. Stimuli were presented to participants in Persian orthography, and speakers read aloud each word. In total, 36 target items contained the uvular phoneme, which occurred in initial, medial and final positions—with 12 words in each position.[1] The canonical form of the uvular consonant is the stop [ɢ]. The uvular phoneme was always in vocalic environments, occurring adjacent to a vowel. Each item was presented twice. Stimuli presentation was randomized.

Data annotation was done in Praat (Boersma & Weenink, 2024). The first author, a native Persian speaker, categorized the manner of the uvular consonant as either stop or approximant. Productions with a closure and a release were classified as a stop, and those with vowel-like qualities, characterized by periodicity in the waveform, were classified as approximants. Productions with low-quality recordings or the inability to classify the target sound into one of the two categories were excluded (10% of all tokens).

Results

In total, productions with the stop allophone comprised 88% of the data, while those with the approximant allophone covered 12% of productions, indicating lower occurrence rates for noncanonical variants. In initial and final positions, canonical variants were observed in more than 97% of productions (Figure 1). In medial position, this category comprised 64% of the productions. Taken together, the results reveal two main patterns. First, the canonical form (i.e., [ɢ]) is more frequent in all positions, confirming that this is a suitable case for testing the representation of low-frequency noncanonical variants. Second, relative occurrence rate of noncanonical variants varies by position: it is more frequent in medial than in initial or final position. As discussed below, this positional difference is critical to testing the predictions of storage of the noncanonical variant.[2]

[1] Items were also tested in a frequency rating task (Likert scale 1 to 10) to assess their frequency relative to frequent and infrequent Persian words which were used to establish the floor and ceiling of ratings. Target item frequency rating medians (8/10) were closer to the rating medians of the frequent (10/10) control items than to the infrequent control items (1/10). In initial (8/10), medial (7/10) and final positions (8/10), lexical frequency medians were similar and close to the medians of the frequent control words.

[2] Productions were examined by item to identify if words with a lexicalized uvular phoneme were produced consistently across participants. Variant production rates for items in each position were similar: in initial and final positions, approximants had a low production rate, and no word was exclusively produced with an approximant. Thus, a single item or a group of items was not responsible for the small number of approximant productions. Similarly, in medial position, all tokens were produced with both variants: not a single item or a group of items was produced with the approximant.

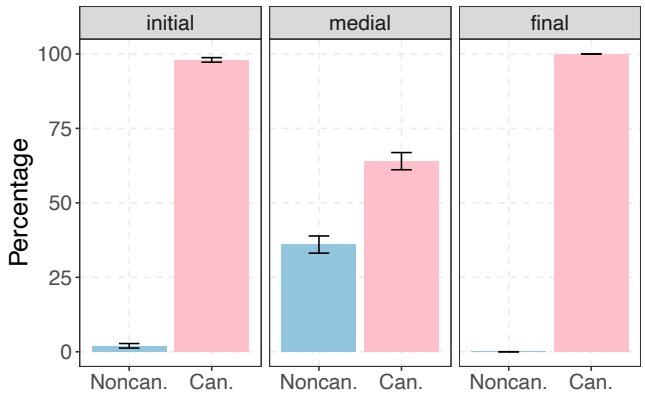

Figure 1: Barplot of the production experiment results. Error bars represent 95% Confidence Interval.

Acceptability rating experiment

In addition to the production experiment, we conducted an acceptability rating task to test the relative frequencies of variants. In the absence of a corpus with phonetic transcriptions, these two experiments served to estimate variant frequencies (Bölte & Coenen, 2002). Specifically, acceptability ratings of variants are reliable estimates for their relative frequency (e.g., Bermel & Knittl, 2012; Brand & Ernestus, 2018; Ranbom & Connine, 2007). Furthermore, Balota et al. (2001) showed that speakers are able to access relative frequency information about lexical entries, and their acceptability ratings can be used as accurate predictions of the relative frequencies of variants observed in corpora.

Methods

Seventy-one Persian speakers (42 f, 29 m, mean age = 30 years) completed the experiment. Stimuli were auditorily presented to participants. On a Likert scale of 1 to 10, they rated how acceptable the pronunciation of each stimulus was and how likely they were to hear that pronunciation in a daily conversation. Stimuli consisted of 12 items in each of three positions (initial, medial, final), and with two variants each (stop/canonical and approximant/noncanonical), for a total of 72 target stimuli. The experiment also included filler items. In total, 36 real words were used. Each filler had two pronunciations: a natural (e.g., [mive] 'fruit') and an unnatural variant with an unacceptable pronunciation that is a nonce word (e.g., [five]). In total, there were 72 fillers in the experiment. Natural and unnatural fillers were included to set the ceiling and floor ratings, respectively and to examine the target variant ratings relative to fillers. All stimuli were randomized.

Results

In each position, the noncanonical form was rated as less frequent and less acceptable than its canonical counterpart (Figure 2). In medial position, the difference between the two variants was smaller than other positions. Overall, rating medians for canonical (10/10) and noncanonical variants (9/10) were closer to rating medians of natural fillers (10/10) than to unnatural fillers (1/10). Even variants with the lowest ratings—word-initial (7/10) and word-final (8/10) approximants—were rated considerably higher than unnatural fillers, indicating these target items are perceived as real words in the language.

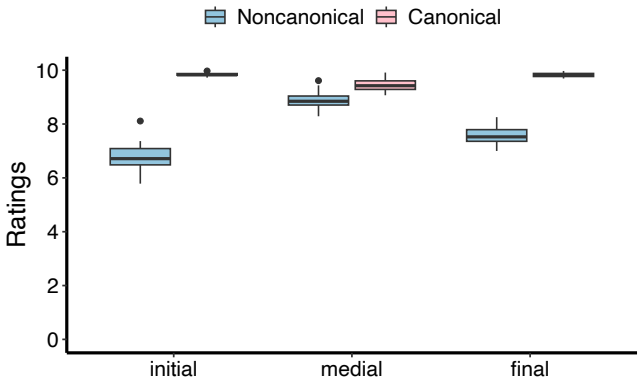

Figure 2: Boxplot of rating experiment results

To examine differences across conditions and positions, ratings were fit to a linear mixed-effects logistic regression (LMER) model with fixed effects Position and Condition (both simple-coded), and random slopes for Position and Variant, along with random intercepts by participant and by item. The statistical results reflected the patterns discussed above: independent of position, noncanonical variants had significantly lower ratings than canonical variants ($\beta = -2.43$, $SE = 0.20$, $z = -12.15$, $p < 0.001$). The difference between the canonical and noncanonical variants in medial position was significantly smaller than their difference in initial ($\beta = -2.96$, $SE = 0.24$, $z = -12.33$, $p < 0.001$) and final position ($\beta = -2.19$, $SE = 0.22$, $z = -9.95$, $p < 0.001$).

These results are similar to the production experiment. In each position, the canonical variant is more acceptable, hence more frequent, than the noncanonical variant, with a smaller difference in medial position, suggesting positional effects on the occurrence rates of variants. As such, the results from the rating and the production experiment converge, confirming test reliability; thus, in the absence of a Spoken Persian corpus, these results can proxy relative frequency and be used to formulate the hypotheses for the representation of multiple pronunciation forms.

Recall that to find evidence for predictions of multiple storage models, there should be an inverse relationship between occurrence frequencies and RTs of variants. Therefore, based on the results of the production and rating experiments, noncanonical variants should overall and in each position elicit slower RTs than canonical variants—given that noncanonical variants overall and within each position had lower ratings than their canonical counterparts. The RT difference between these two variants should be smaller in medial position than in other positions, suggesting a positional effect of variants. Obtaining such an inverse relationship between occurrence

frequencies and RTs would support predictions of multiple storage systems. By contrast, single storage models do not predict positional effects; instead, RT differences between variants should be constant across all positions, with the canonical variant eliciting consistently faster RTs than the noncanonical variant in each position. That is, if the canonical variant is the only represented form, there should be a constant RT difference between variants across all positions, again suggesting no positional effects.

Lexical decision experiment

In this experiment, participants listened to the experimental items and decided whether they were real or nonce words via keyboard button press. Before presenting each stimulus, a fixation point ("+") appeared in the middle of the screen for 250 ms. Then, the audio stimulus was presented. RTs and accuracy were measured from the stimulus offset. The intertrial interval was 1000 ms.

Methods

This experiment involved the same participants, stimuli and variants as the rating experiment where 75% of stimuli were real words. To avoid bias in the lexical decision task due to this imbalance, 72 additional nonce word fillers were added to equalize the proportion of real and nonce words.

Before examining results, trial RTs greater than 10 seconds were removed (1% of all trials). Following Baayen (2008), Goodwin Davies and Embick (2020), and Taghipour and Monahan (2020), trials with RTs ±2.5 SD from each individual participant's average RTs were also removed (3.3% of all trials). Responses for stimuli that were judged as "nonword" were removed (24% of all trials). Participants considered 35% of noncanonical and 7% of canonical variants as nonwords. Most nonword responses were for noncanonical variants with word-initial and word-final approximants. No variant or item was consistently deemed a "nonword" by participants; rather, all items received some nonword judgments, but with varying proportions. Participants had overall similar judgments patterns, and no specific participant judged all items as nonwords, suggesting no item- or participant-specific effects. The higher proportion of nonword judgments for approximants than stops could stem from their low occurrence rates, especially at word edges; with less exposure to approximants, they are more likely to be perceived as nonwords than the more frequent stops.

Results

In each position, the noncanonical variant had slower RTs than the canonical variant (Figure 3). In medial position, the RT difference of the two variants was smaller than their difference in other positions. RTs were fit to a LMER model. The primary questions were whether there was an effect of canonical versus noncanonical variants, and whether this effect differed across position, specifically comparing medial versus initial and medial versus final positions. Overall, the noncanonical variant had significantly slower RTs than the canonical variant ($\beta = 138.54$, SE = 23.18, $z = 5.97$, $p < 0.001$). The difference between canonical and noncanonical variants in medial position was smaller than their difference in initial ($\beta = 93.53$, SE = 32.61, $z = 2.86$, $p = 0.005$) and final positions ($\beta = 119.11$, SE = 29.44, $z = 4.04$, $p < 0.001$).

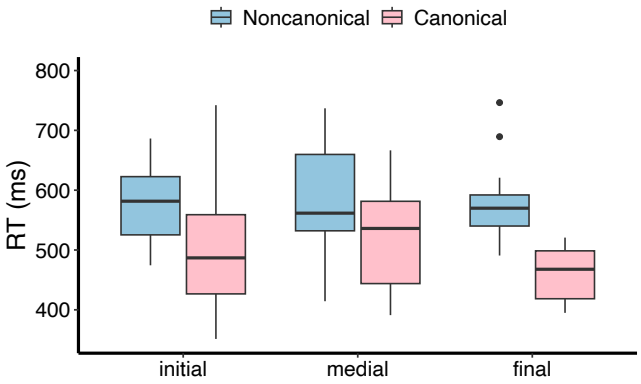

Figure 3: Boxplot of lexical decision experiment results

In summary, results of production, rating and lexical decision experiments exhibited the predicted inverse relationship between occurrence frequencies and RTs independent of position, as well as across positions. These positional effects are consistent with predictions of multiple storage accounts.

Discussion

With the goal of investigating predictions of multiple storage models, we examined whether low frequency noncanonical forms can be represented in the lexicon. Results revealed that, contrary to predictions of single storage models, a noncanonical variant can be stored without being overwhelmingly frequent. Specifically, a low frequency noncanonical variant can be lexically stored, the same way a canonical variant is lexically stored. Together, the observed positional effects across these experiments and using low-frequency noncanonical variants provide robust evidence in support of the representation of such noncanonical variants and confirm the predictions of multiple storage models.

Such results have implications for models of speech perception. According to single storage accounts, processing and recognition of noncanonical variants should take longer due to the mismatch between the features of the input and those of the canonical representation (Bölte & Coenen, 2002); single storage model predictions suggest that when a canonical variant is the input, it should be processed faster because features of the input and the represented form match, leading to faster recognition of the input. Therefore, in the context of speech processing, compared to their noncanonical counterparts, canonical forms have a privileged status because only the canonical variant is stored, and processing of all pronunciation variants depends on the degree of similarity to their canonical counterpart. However, based on the current evidence for the separate representation of noncanonical variants, it can be postulated that when a noncanonical variant is

the input, it is not necessarily compared to its canonical form; instead, the noncanonical input is directly mapped to its representation, highlighting the storage of noncanonical variants in addition to their canonical counterparts. This perspective suggests that there is no difference between a canonical and a noncanonical variant during spoken word recognition; as such, compared to their noncanonical counterparts, canonical forms do not seem to have a privileged status because within the multiple storage framework, variants are processed based on their occurrence frequencies, and not their canonicity (see Bürki et al., 2010; cf. Ernestus & Baayen, 2007). Thus, if processing of a noncanonical variant, like the one examined here, takes longer than its canonical counterpart, it is not necessarily due to the feature mismatch between the noncanonical input and its canonical representation; rather, it may be due to the lower frequency of the noncanonical variant. If noncanonical variants had higher frequencies, they would exhibit faster RTs than their canonical counterparts. Such a result underscores the role of occurrence frequency in modulating processing and representation of phonetic variants.

These observations and findings also align with those of previous studies that examine the representation of noncanonical variants. For instance, Brand and Ernestus (2018) showed that beginner learners of French had not yet represented noncanonical variants due to low occurrence rates and limited exposure; in contrast, advanced learners and native speakers represented these variants, because of their higher exposure and greater occurrence rates of such forms—emphasizing the role of occurrence rate and exposure on the representation of noncanonical variants. Single storage models may encounter challenges explaining such outcomes. According to these models, canonicity of a variant, and not occurrence frequency, determines its processing time; therefore, within the single storage models framework, one would predict that beginners, advanced learners and native speakers should process noncanonical variants similarly; however, Brand and Ernestus obtained different results, showing clear effects of occurrence frequency on the processing time of variants. These results suggest that one critical factor, if not the most important one, is variants relative frequency in their representation, and not necessarily their canonicity.

This perspective within multiple storage accounts, postulating the effect of occurrence frequency of variants on their recognition, offers a dynamic and nuanced view of the lexicon with respect to variant representation compared to single storage models. Overall, single storage accounts presume a fixed and static view of lexicon due to assuming only a single canonical representation per word and disregarding the role of variant frequency (e.g., Caramazza, 1997; Connine et al., 1993; Dell, 1990; McClelland & Elman, 1986; Norris, 1994; Stevens, 2002). In contrast, results from proponents of multiple storage models (e.g., Bürki et al. 2010; Johnson, 2004; Pierrehumbert, 2001; 2002) and the findings of the present research imply a more dynamic perspective, as these studies show that noncanonical forms can also be lexically represented. This perspective can also account for the representation of noncanonical variants in the context of learning a foreign language, as in Brand and Ernestus (2018). Their study implies that, like advanced learners, with sufficient exposure and occurrence frequency, beginner learners could develop representations for noncanonical variants. This observation indicates that new noncanonical representations can be incorporated into the lexicon, suggesting a dynamic perspective on the lexicon, which can be continually updated. This dynamic view aligns with the assumption that lexicon is not a static repository but rather a flexible and evolving system that can adapt as speakers encounter and process multiple pronunciation word forms over time (see Pierrehumbert, 2016). In this sense, the lexicon is a system in which variant representations are continuously shaped depending on the input and usage patterns, reinforcing the importance of frequency and exposure in the representation of canonical and, particularly, noncanonical forms in the lexicon.

To conclude, via testing the representation of low-frequency noncanonical forms, this research examined the predictions of multiple storage systems. Finding evidence for the separate storage of noncanonical forms confirms the predictions of these models, and crucially, are not consistent with predictions of any proposed single storage model. These results suggests that a noncanonical variant need not be high-frequency to be lexically represented. Additionally, extending research in the domain of lexical representations to the storage of Persian uvular variants expands the scope of the predictions of multiple storage accounts to understudied phenomena and languages, emphasizing the need to explore lesser-known patterns to broaden the generalizability of speech representation models.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388-407. https://doi.org/10.3758/s13428-019-01237-x

Ariyaee, K. (2021). Acoustics of Persian uvular lenition. *Unpublished Generals Paper.* University of Toronto.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge University Press.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition, 29*, 639-647. https://doi.org/10.3758/BF03200465

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*(1), 92-111. https://doi.org/10.1016/j.jml.2008.06.003

Bermel, N., & Knittl, L. (2012). Corpus frequency and acceptability judgments: A study of morphosyntactic variants in Czech. *Corpus Linguistics and Linguistic Theory, 8*(2), 241-275. https://doi.org/10.1515/cllt-2012-0010

Boersma, P., & Weenink, D. (2024). Praat: doing phonetics by computer [Computer program] (version 6.4.20).

Bölte, J., & Coenen, E. (2002). Is phonological information mapped onto semantic information in a one-to-one manner? *Brain and Language, 81*(1-3), 384-397. https://psycnet.apa.org/doi/10.1006/brln.2001.2532

Blumstein, S. E. (2004). Phonetic category structure and its influence on lexical processing. In A. Agwuele et al. (Eds.), *Proceedings of the Texas Linguistic Society* (pp. 17-25). Somerville, MA: Cascadilla Press.

Brand, S., & Ernestus, M. (2018). Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: Evidence from native listeners and learners of French. *Quarterly Journal of Experimental Psychology, 71*(5), 1240-1259.

Bürki, A. (2018). Variation in the speech signal as a window into the cognitive architecture of language production. *Psychonomic Bulletin Review, 25*(6), 1973-2004. https://doi.org/10.3758/s13423-017-1423-4

Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language, 62*(4), 421-437. https://doi.org/10.1016/j.jml.2010.01.002

Bürki, A., & Gaskell, M. G. (2012). Lexical representation of schwa words: Two mackerels, but only one salami. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 617-63. https://doi.org/10.1037/a0026167

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology, 14*(1), 177-208. https://doi.org/10.1080/026432997381664

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*(2), 193-210. https://doi.org/10.1006/jmla.1993.1011

Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*, 313-349. https://doi.org/10.1080/01690969008407066

Dufour, S., Fléchard, L., & Nguyen, N. (2022). Phonological abstraction before lexical access: New evidence from rime priming. *Quarterly Journal of Experimental Psychology, 75*(12), 2211-2218. https://doi.org/10.1177/17470218221077917

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*, 27-41. https://doi.org/10.1016/j.lingua.2012.12.006

Ernestus, M., & Baayen, R. H. (2007). The comprehension of acoustically reduced morphologically complex words: The roles of deletion, duration and frequency of occurrence. In *Proceedings of ICPHS* (pp. 773-776). Saarbrücken, Germany.

Fougeron, C., Goldman, J.-P., & Frauenfelder, U. H. (2001). Liaison and schwa deletion in French: An effect of lexical frequency and competition? *Proceedings of Euro speech 2001* (pp. 639-642). Aalborg, Danemark.

Goodwin Davies, A., & Embick, D. (2020). The representation of plural inflectional affixes in English: Evidence from priming in an auditory lexical decision task. *Language, Cognition and Neuroscience, 35*(3), 393–401. https://doi.org/10.1080/23273798.2019.1684528

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.) Spontaneous Speech: Data and Analysis. *Proceedings of the 1st Session of the 10th International Symposium* (pp. 29-54). Tokyo, Japan: The National International Institute for Japanese Language.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. The MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75. https://doi.org/10.1017/s0140525x99001776

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics, 62*(3), 615-625. https://doi.org/10.3758/BF03212113

Marslen-Wilson, W., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1376- 1392. https://doi.org/10.1037/0096-1523.22.6.1376

McClelland, J. L., & Elman, J. L. (1986). *The TRACE model of speech perception. Cognitive Psychology, 18*(1), 1-86. https://doi.org/10.1016/0010-0285(86)90015-0

McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. In A. S. Meyer & N. O. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*. Mouton de Gruyter.

Mohammadi, A. N. (2019). Corpus of Conversational Persian Transcripts LDC2019T11. Web Download. Philadelphia: Linguistic Data Consortium. https://doi.org/10.35111/qxep-jt47

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234. https://doi.org/https://doi.org/10.1016/0010-0277(94)90043-4

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. John Benjamins Publishing Company. https://doi.org/10.1075/tsl.45.08pie

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 101-140). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110197105.1.101

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics, 2*(1), 33-52. https://doi.org/10.1146/annurev-linguistics-030514-125050

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language, 61*(1), 19-36. https://doi.org/10.1016/j.jml.2009.02.005

Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language, 57*(2), 273-298. https://doi.org/10.1016/j.jml.2007.04.001

Samuel, A. G., & Larraza, S. (2015). Does listening to non-native speech impair speech perception? *Journal of Memory and Language, 81*, 51-71. https://doi.org/10.1016/j.jml.2015.01.003

Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A., & Deco, G. (2009). Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence. *Journal of Cognitive Neuroscience, 21*(12), 2343-2357. https://doi.org/10.1162/jocn.2008.21152

Soo, R., & Babel, M. (2024). Regular exposure facilitates dual-mapping of Cantonese pronunciation variants. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique, 69*(1), 99-117. doi:10.1017/cnj.2024.3

Spinelli, E., & Gros-Balthazard, F. (2007). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition, 104*(2), 397-406. https://doi.org/10.1016/j.cognition.2006.07.002

Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America, 111*(4), 1872-1891. https://doi.org/10.1121/1.1458026

Sumner, M. (2013). A phonetic explanation of pronunciation variant effects. *The Journal of the Acoustical Society of America, 134*(1), EL26-EL32. https://doi.org/10.1121/1.4807432

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487-501. https://doi.org/10.1016/j.jml.2009.01.001

Sumner, M., Kim, S. K., King, E., & McGowan, K. B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology, 4*, 66115. https://doi.org/10.3389/fpsyg.2013.01015

Taghipour, S., & Monahan, P. J. (2020). Early Morphological Decomposition: The Effect of Derivational Gaps. *Proceedings of the 38th Annual Conference of the Canadian Linguistic Association*.