Patterns of Misperception of Arabic Consonants

Chelsea Sanker

7 December 2014

1 Introduction

1.1 Misperception Experiments

- Previous work on misperception has shown a range of influences such as native language (e.g. Cutler et al. 2004), position in the syllable (e.g. Wang and Bilger 1973), and types of masking noise (e.g. van Phatak, Lovitt, and Allen 2008).
- Can patterns of misperception among oral consonants be extrapolated to guttural consonants (uvulars, pharyngeals, and glottals)?
- Data on consonant discrimination in Arabic (e.g. Kishon-Rabin and Rosenhouse 2000) and the phonological characteristics of gutturals (cf. McCarthy 1994) suggest that guttural consonants are likely to have their own misperception patterns; this study investigates those patterns.
- While there have been studies on misperception of illicit structures (e.g. Dupoux et al. 1999), there have not, to my knowledge, been studies on the patterns of misperception that are present in identification of licit structures; this study investigates such patterns.

1.2 Features of Guttural Consonants

- There are many behaviors which guttural consonants share, such as causing retraction and lowering on neighboring vowels (McCarthy 1994); which gestures do they share? Which features do listeners use to identify them?
- There is disagreement about the features which characterize guttural consonants; patterns of confusions can help elucidate the salient acoustic features shared by consonants (Chang, Plauché, and Ohala 2001).

1.3 Historical Implications of Misperception Patterns

- There are many parallels between misperceptions of sounds and sequences in laboratory experiments and attested historical changes (Blevins and Garrett 2004); sound change can result when people insufficiently or inaccurately account for contextual influence (Ohala 2003).
- Some studies have suggested that pharyngeal consonants can be lost via vocalization, e.g. in Kalispel, though some apparent cases of vocalization seem to be due to non-native transcribers mishearing these consonants (Bessell 2002)
- Pharyngeals are often reconstructed in Proto-Indo-European, for segments with vocalic outcomes in many daughter languages.

2 Experiment Design

2.1 Participants

There were two groups of participants:

1. Native speakers of Levantine Arabic (n = 3, ages 20-35) who spent the majority of their childhood in a country where Arabic was the primary language.
2. Native speakers of English (n = 6, ages 20-27) who have been studying Arabic for 2-3 years.

2.2 Stimuli

Stimuli were nonce words produced by native speakers of Arabic in Modern Standard Arabic (MSA).

Syllable shapes were CV, ʔVC, ʔVRC, ʔVRV, and ʔVRVC, where R is each of the liquids (l, r). Consonants were produced in the context of each of the contrastive Arabic vowel qualities: /i/, /a/, /u/.

2.3 Task

- In the main task, listeners saw an array of buttons of nonce words written in the Arabic script and listened to the stimuli (their order randomized within blocks of the same vowel), then clicked the button corresponding to the word which they thought they heard.
- Masking noise of perceptual loudness equal to the stimulus was played simultaneously with the stimuli, in order to increase the difficulty of the task and elicit a higher number of errors. The two noise types were: noise with intensity weighted towards lower frequencies (pink noise) or towards higher frequencies (blue noise).

3 Data and Discussion

In addition to differences in the average accuracy in each group, there were differences in the sounds which were most frequently confused and in the directions of confusion in each group, as well as effects of phonological environment, the consonant's place of articulation, and the type of masking noise.

3.1 Structure Confusions

Table 1: Native Listeners' Structure Identifications (as percents)

stimulus \ item selected	ʔVC	ʔVRC	ʔVRV	ʔVRVC
ʔVC	96	3	1	0
ʔVRC	6	82	2	10
ʔVRV	4	3	75	18
ʔVRVC	0	9	8	83

Table 2: Non-native Listeners' Structure Identifications (as percents)

	ʔVC	ʔVRC	ʔVRV	ʔVRVC
ʔVC	92	5	1	2
ʔVRC	19	62	3	15
ʔVRV	4	8	69	19
ʔVRVC	2	24	12	62

- Although the rate of misperceptions of structure was lower in the native listener group, the patterns of misperception were similar in both groups.
  - Non-native listeners made more structural errors that involved the deletion of a segment, while native listeners made a similar number of errors that involved the insertion or deletion of a segment.
- There was a difference between structural accuracy for words with guttural consonants or without them, which interacted with language background.
  - Native Arabic listeners identified structure with 92% accuracy for words without gutturals and 91% for words with gutturals
  - Non-native listeners identified them with 88% and 79% accuracy, respectively.
  - There was a high level of confusion with vowels between consonants and with certain word-final consonants after vowels: [h] and [ʔ] were often identified as present when they were not, and the guttural consonants as well as the alveolar stops were prone to deletion in biconsonantal structures. Cf. deletion of final glottal consonants in spoken Arabic.

3.2 Segment Confusions

Native Vowel and Consonant Confusions

There are cases in both groups of listeners in which a consonant was identified as a vowel or a vowel as a consonant, though these identifications were much less common than identifications involving the addition or loss of a consonant or vowel.

Table 3: Native Listeners' Segment Confusions Involving Vowels (as percents)

	a	i	u	j	w	ʕ	ʔ	h	∅	other
a	**96**	-	-	-	-	-	-	-	-	-
i	-	**88**	-	4	-	-	-	-	8	-
u	-	-	**96**	-	-	-	-	4	-	-
j	-	2	-	**89**	-	-	-	10	-	-
w	-	-	3	-	**98**	-	-	-	-	-
ʕ	1	-	-	-	-	**70**	13	3	4	10
ʔ	3	-	-	2	1	4	**78**	4	3	7
h	2	-	-	-	1	2	8	**65**	3	20

- In the native listener group, most of the common misperceptions were not strongly paralleled by misperceptions in the opposite direction, e.g. there were 3 times as many cases of [t] identified as [k] than [k] identified as [t].
- The most common misperception error in the native listener group was hearing the voiceless uvularized alveolar stop [tˤ] as the voiceless uvular stop [q].

Non-native Vowel and Consonant Confusions

Table 4: Non-native Listeners' Segment Confusions Involving Vowels (as percents)

	a	i	u	j	l	t	d	k	x	ʁ	ħ	ʕ	ʔ	h	∅	other
a	**84**	-	-	-	-	-	-	-	-	5	2	2	7	-	-	-
i	2	**80**	2	6	-	-	-	2	-	-	-	2	-	4	-	-
u	-	-	**71**	-	-	2	-	-	2	-	-	2	-	21	2	-
j	1	1	-	**77**	6	1	-	-	-	1	2	3	3	3	-	1
l	1	-	-	2	**66**	-	10	1	-	8	1	2	-	-	-	6
t	-	-	-	-	-	**59**	-	8	3	-	5	-	3	1	6	14
d	-	-	1	2	-	5	**62**	-	2	3	-	-	1	-	10	15
k	1	-	-	1	-	9	-	**44**	5	2	5	2	3	1	3	24
x	-	-	-	-	-	2	-	3	**61**	4	13	2	3	3	3	5
ʁ	-	-	2	-	1	-	1	-	2	**59**	1	4	3	6	4	7
ħ	1	-	-	3	-	2	1	1	4	1	**42**	5	3	19	10	6
ʕ	2	1	-	-	-	-	-	2	3	1	6	**38**	27	5	8	6
ʔ	2	-	-	1	-	2	-	3	1	2	6	11	**52**	2	14	2
h	-	-	-	-	-	1	-	1	4	5	32	5	10	**24**	12	5

- The most common confusions in the non-native group of listeners were between uvularized consonants and their non-uvularized counterparts (e.g. [s] and [sˤ]), and between pharyngeals and glottals (e.g. [h] and [ħ]).

See Tables 7 and 8 in the appendix for the full consonant confusion matrices.

Influences on segment accuracy

- The accuracy of consonant identification was influenced by the phonological environment in both groups of listeners:
  - Native listeners identified consonants with 84% accuracy after vowels and 78% accuracy after liquids, while non-native listeners identified them with 66% accuracy after vowels and 52% accuracy after consonants.
  - In both groups of listeners, consonants were more accurately identified in codas than in onsets.
- The accuracy of segment identification differed by the region of the consonantal constriction.
  - Non-native listeners had significantly lower accuracy in identifying guttural consonants (45% correct) than consonants with other places of articulation (68% correct).
  - For native listeners, this difference was not present: guttural consonants were identified with 79% accuracy, while other consonants were identified with 80% accuracy.
- The average response time was slightly longer for native speakers (mean = 2.6 s) than non-native speakers (mean = 2.4 s). The response times for sounds in each category also differed.
  - For native speakers, the mean response time for guttural consonants was 3.7 s, and for non-guttural consonants was 4 s.
  - For non-native speakers, the mean response time for guttural consonants was 3.5 s and for non-guttural consonants was 2.9 s.
- There was an effect of noise type on accuracy of identification; sounds were identified more accurately in higher frequency noise than in lower frequency noise.
  - Native listeners identified consonants with 83% accuracy in high frequency noise and 68% accuracy in low frequency noise.
  - Non-native listeners identified consonants with 64% accuracy in high frequency noise and 51% accuracy in low frequency noise.
  - This effect interacted with place of articulation; guttural consonants seemed to suffer more interference from pink noise than non-guttural consonants did.

Table 5: Accuracy with Noise Type and Consonant Region for Native Listeners (as percents)

	blue	pink
guttural	83	66
non-guttural	82	72

Table 6: Accuracy with Noise Type and Consonant Region for Non-native Listeners (as percents)

	blue	pink
guttural	55	40
non-guttural	72	64

4 Conclusions

- There are differences in the behavior of guttural and non-guttural consonants, so one should be cautious about extrapolating from oral consonant data to guttural consonants.
- While change from guttural consonants into vowels is not impossible, a two-step change of vowel insertion followed by consonant loss seems more likely, particularly because both of these changes are relatively common misperceptions for guttural consonants, particularly in certain conditions (e.g. low frequency noise and non-native listeners).
- The primary cues for identification of consonants and in particular guttural consonants are contained in the lower frequencies.
- There are patterns of directionality in misperception of guttural consonants, which do not consistently match up with attested or commonly reconstructed sound changes.

References

[1] Bessell, Nicola Jane 1992. Towards a Phonetic and Phonological Typology of Post-Velar Articulation. Ph.D., University of British Columbia.

[2] Blevins, Juliette, and Andrew Garrett. 2004. The evolution of metathesis. In *Phonetically Based Phonology*. Ed. Bruce Hayes, Robert Kirchner, and Donca Steriade. Cambridge: Cambridge University Press.117-156.

[3] Chang, Steve, Madelain Plauché, and John Ohala. 2001. Markedness and Consonant Confusion Asymmetries.? In textitThe Role of Speech Perception in Phonology. Eds. B. Hume and K. Johnson. 79-101. San Diego: Academic Press.

[4] Cutler, Anne, Andrea Weber, Roel Smits, and Nicole Cooper. 2004. Patterns of English Phoneme Confusions by Native and Non-Native Listeners. *Journal of the Acoustical Society of America* 116.6. 3668-3678.

[5] Dubno, Judy, and Harry Levitt. 1981. Predicting Consonant Confusions form Acoustic Analysis. *Journal of the Acoustical Society of America* 69.1. 249-261.

[6] Dupoux, Emmanuel, Erika Parlato, Sonia Frota, Yuki Hirose, and Sharon Peperkamp. 2011. Where do Illusory Vowels Come From? *Journal of Memory and Language* 64.199-210.

[7] Kishon-Rabin, Liat and Judith Rosenhouse. 2000. Speech perception test for Arabic-speaking children. *Audiology* 39:5. 269-277.

[8] McCarthy, John. 1994. The phonetics and phonology of pharyngeals. In Patricia Keating, ed., *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, 191-233. Cambridge: Cambridge University Press.

[9] Ohala, John. 2003. Phonetics and Historical Phonology. In *The Handbook of Historical Linguistics*. Ed. Brian D. Joseph and Richard D. Janda. Oxford: Blackwell Publishing.669-686.

[10] van Phatak, Sandeep, and Jont B. Allen. 2007. Consonant and Vowel Confusions in Speech-Weighted Noise. *Journal of the Acoustical Society of America* 121.4. 2312-2326.

[11] Wang, Marilyn and Robert Bilger. 1973. Consonant Confusions in Noise: A study of Perceptual Features. *Journal of the Acoustical Society of America* 54.5. 1248-1266.

Table 7: Native Listeners' Segment Identifications (as percents)

	j	w	r	l	t	tˤ	d	dˤ	s	sˤ	z	ʤ	k	q	x	ʁ	ħ	ʕ	ʔ	h	∅
j	**89**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	-
w	-	**98**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
r	-	2	**89**	-	-	-	-	-	-	-	-	-	-	2	-	3	-	2	3	-	-
l	2	3	-	**69**	-	-	9	-	-	-	2	-	-	-	-	10	-	-	2	1	-
t	-	-	-	-	**68**	1	1	-	-	-	-	-	14	3	1	2	-	1	3	3	5
tˤ	-	2	-	-	2	**66**	-	4	-	-	-	-	-	23	2	-	-	-	2	-	-
d	-	-	-	-	2	-	**86**	1	-	-	1	-	1	-	-	3	-	-	1	-	6
dˤ	4	-	-	-	2	-	4	**87**	-	-	2	-	-	-	-	-	-	2	-	-	-
s	-	-	-	-	4	-	-	-	**83**	3	-	-	1	1	4	1	-	1	-	2	1
sˤ	-	-	-	-	-	-	-	-	10	**90**	-	-	-	-	-	-	-	-	-	-	-
z	-	-	-	-	-	-	3	-	-	-	**94**	-	-	-	-	3	-	-	-	-	-
ʤ	-	-	-	-	-	-	11	-	-	-	-	**83**	-	-	3	3	-	-	-	-	-
k	-	-	-	-	5	4	1	-	-	-	-	-	**68**	8	4	2	-	-	4	2	2
q	1	1	-	-	-	3	-	-	-	-	-	-	2	**88**	3	-	-	-	3	-	1
x	-	-	-	-	1	1	-	-	-	-	-	1	-	1	**87**	1	2	-	-	5	2
ʁ	-	5	3	-	-	-	-	2	-	-	-	-	-	4	-	**80**	-	1	-	2	5
ħ	1	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	**84**	3	1	5	4
ʕ	-	-	-	-	-	-	-	-	-	-	-	-	1	2	2	2	3	**70**	13	3	4
ʔ	2	1	-	1	1	-	2	-	-	-	-	-	-	1	-	1	1	4	**78**	4	3
h	-	-	-	-	1	-	2	1	-	-	-	-	-	2	2	1	14	-	8	**65**	3
∅	-	-	-	-	-	1	-	-	-	-	-	-	1	1	1	1	1	-	4	5	**87**

Table 8: Non-native Listeners' Segment Identifications (as percents)

	j	w	r	l	t	tˤ	d	dˤ	s	sˤ	z	ʤ	k	q	x	ʁ	ħ	ʕ	ʔ	h	∅
j	**77**	6	-	1	-	-	-	-	-	-	-	1	-	-	1	2	2	3	3	3	-
w	3	**94**	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-
r	-	-	**76**	2	-	-	-	2	-	-	-	-	1	1	-	12	-	-	1	1	-
l	2	4	1	**66**	-	-	10	2	-	-	-	-	1	-	-	8	1	2	-	-	-
t	-	-	-	-	**59**	11	-	-	-	-	-	-	8	3	3	-	5	-	3	1	6
tˤ	-	4	1	-	11	**38**	3	7	-	-	1	-	1	14	1	6	6	3	1	1	-
d	2	-	-	-	5	2	**62**	12	-	-	-	1	-	-	2	3	-	-	1	-	10
dˤ	-	-	1	1	-	6	17	**61**	-	-	1	-	1	-	-	7	-	1	1	-	-
s	-	-	-	-	6	-	-	-	**73**	19	-	-	-	-	-	-	-	-	-	-	-
sˤ	-	-	-	-	3	2	-	-	40	**55**	-	-	-	-	-	-	-	-	-	-	-
z		-	-	-	-	-	1	1	-	-	**96**	-	-	-	-	1	-	-	-	-	-
ʤ	-	-	-	-	1	-	4	6	-	-	-	**89**	-	-	-	-	-	-	-	-	-
k	1	-	-	-	9	1	-	-	-	-	-	-	**44**	23	5	2	5	2	3	1	3
q	-	1	-	-	3	2	1	-	-	-	-	-	12	**54**	6	8	3	2	3	-	3
x	-	-	-	-	2	1	-	-	1	-	-	-	3	3	**61**	4	13	2	3	3	3
ʁ	-	6	7	1	-	1	1	-	-	-	1	-	-	2	2	**59**	1	4	3	6	4
ħ	3	1	1	-	2	-	1	-	-	-	-	-	1	4	4	1	**42**	5	3	19	10
ʕ	-	-	1	-	-	-	-	-	-	-	-	-	2	5	3	1	6	**38**	27	5	8
ʔ	1	-	-	-	2	-	-	-	-	-	-	-	3	2	1	2	6	11	**52**	2	14
h	-	2	-	-	1	1	-	-	-	-	-	-	1	2	4	5	32	5	10	**24**	12
∅	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-	1	1	1	4	5	**77**