GREEK CONSONANT CONFUSIONS BY NATIVE LISTENERS IN QUIET AND NOISE

Angelos Lengeris & Katerina Nicolaidis
Aristotle University of Thessaloniki
lengeris@enl.auth.gr, knicol@enl.auth.gr

Abstract

This study examined the perception of Greek consonants by native listeners. Consonants were embedded in VCV tokens and were presented for identification in three listening conditions: in quiet and in two types of noise, a competing talker and an 8-speaker babble (signal-to-noise ratios at -6 dB and -2 dB respectively). Listeners' identification in quiet was significantly higher than in noise. The 8-speaker babble had a greater deteriorating effect in listeners' identification than the competing speaker. Across listening conditions, voicing was more resistant to noise than manner of articulation which was in turn more resistant to noise than place of articulation.

Keywords: Greek consonants, identification, noise

1. Introduction

Listeners are faced with the task of understanding speech in various listening conditions on a daily basis. Under ideal conditions, phoneme recognition is relatively easy (although not perfect especially for some phonemes) but when dealing with degraded speech, e.g. when listening speech in the presence of background noise phoneme confusions are much more common.

Although listeners generally manage to overcome such difficulties using other types of information (e.g. contextual, visual) the patterns of confusions in imperfect conditions can benefit our understanding of speech perception. Previous research has shown for instance that onset consonants are more accurately identified than coda consonants (e.g. Benkí 2003, Redford & Diehl 1999, Wang & Bilger 1973) and that place of articulation errors are more frequent than voicing and manner of articulation errors (e.g. Miller & Nicely 1955, Wang & Bilger 1973, Phatak & Allen 2007, Phatak, Lovitt & Allen 2008). In addition, across different types of noise some consonants are consistently confused with others (e.g. /f/ is confused with /θ/ in English, see Miller & Nicely 1955, Redford & Diehl 1999).

While the number of studies concerned with the phonetics of Modern Greek is growing, research on the perception of Greek consonants by native listeners either in quiet or in noise is still lacking (for a comprehensive review of research on Greek phonetics, see Arvaniti 2007). This paper is part of a larger project investigating the perception and production of English (second-language) and Greek (native) consonants and vowels by Greek speakers. Here we present results on the identification of Greek consonants in quiet and in the presence of two types of noise, a competing talker at a signal-to-noise (SNR) of -6dB and an 8-speaker babble at an SNR of -2dB. Our goal is to provide a large source of data on consonant confusions by native listeners under different listening conditions covering the complete

In G. Kotzoglou et al. (eds), 2014, *Selected Papers of the 11th International Conference on Greek Linguistics*, 866-873. Rhodes: University of the Aegean.

phoneme inventory of Modern Greek, using non-word materials that eliminate any word-frequency effects.

2. Method

2.1 Participants

The participants were 20 female speakers of Greek with a mean age of 19.8 years old (aged 19-20), all students at the Aristotle University of Thessaloniki Department of Linguistics. All participants reported normal hearing and no language impairment.

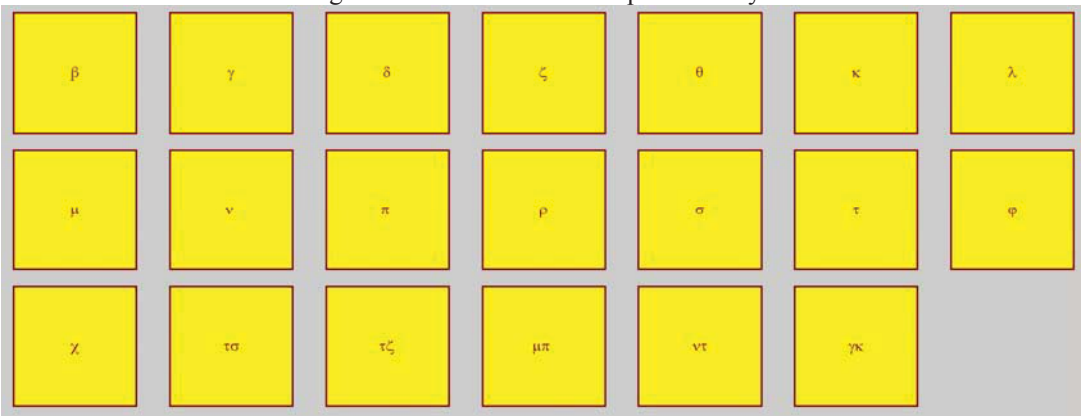
2.2 Perceptual stimuli

The perceptual stimuli consisted of VCV tokens containing the stops /p/, /b/, /t/, /d/, /k/ and /g/, the fricatives /f/, v/, /θ/, /ð/, /s/, z/, x/ and /ɣ/, the nasals /m/ and /n/, the liquids /l/ and /ɾ/ and the affricates /ts/ and /dz/. All 20 consonants were recorded in the context of 3 vowels (/i/, /a/, /u/) in 9 possible combinations. Each VCV token was spoken with stress on the first or second syllable (e.g. /ˈiku/, /aˈsi/, /ˈanu/) by four native speakers of Standard Modern Greek (2 female and 2 male). The VCV tokens were presented in quiet (QUIET) and in the presence of two noise types, a competing speaker (COMP) and an 8-speaker babble (BABBLE). Maskers were taken from a freely available corpus containing, among others, sentences spoken by British English speakers (Cooke & Scharenborg 2008). In COMP, stimuli were presented to participants at a signal-to-noise ratio (SNR) of -6 dB and in BABBLE at an SNR of -2 dB. Each of three test conditions (QUIET, COMP, BABBLE) contained two instances of each of the 20 consonants spoken from the four Standard Modern Greek speakers for a total of 160 VCV items per test condition (grand total of 3 × 160 = 480 items).

2.3 Procedure

Participants were tested individually in the Phonetics Laboratory of the School of English, Aristotle University using a laptop computer and high-quality headphones (Sennheiser HD 280 Professional). Following each VCV item presentation (designed using PRAAT software, see Boersma & Weenink 2008), participants indicated which consonant they heard by clicking on a computer screen one of 20 consonant options written in Greek orthography (Fig. 1) using a mouse. QUIET was always presented first, followed by the two noise conditions (half of the times COMP followed QUIET and half of the times BABBLE followed QUIET). A practice task with 20 VCV tokens for each condition preceded testing to familiarize participants with the procedure.

Figure 1: Screenshot of the experiment layout.



3. Results

Figure 2 shows mean percent correct identification scores pooled over Greek consonants and speakers in each test condition. Identification scores in QUIET were near ceiling (96% correct). Identification scores in COMP were much lower (69% correct) than scores in QUIET but higher than scores in BABBLE (48% correct). A repeated-measures ANOVA on percentages of correct responses confirmed a significant effect of Test condition on identification scores [$F(2,38) = 317.35$, $p<0.001$]. Pairwise comparisons showed that participants achieved higher identification scores in QUIET than in COMP and BABBLE and that their identification scores in COMP were higher than their scores in BABBLE $p<0.001$.

Figure 2: Mean percent correct identification of Greek consonants in each test condition.

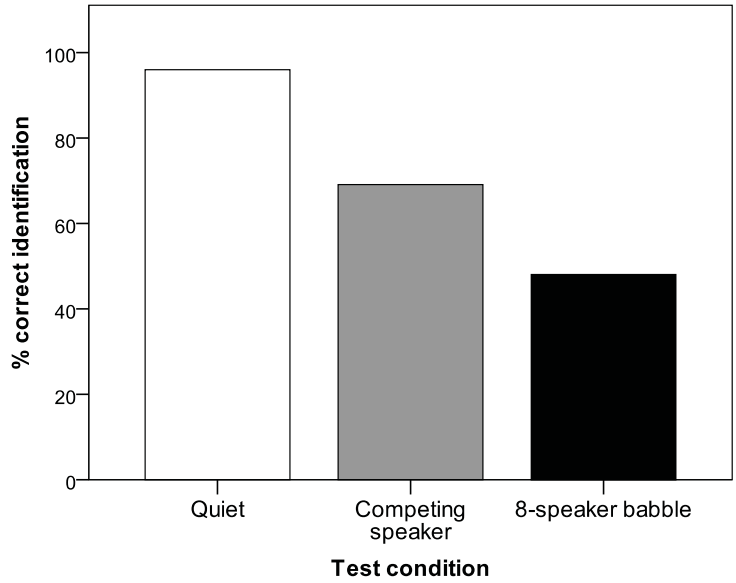


Figure 3 shows percent correct identification for each Greek consonant in each test condition. A repeated-measures ANOVA on percentages of correct responses with Test condition and Consonant as factors showed a significant main effect of

Figure 3: Percent correct identification for each of the 20 Greek consonants in each test condition.

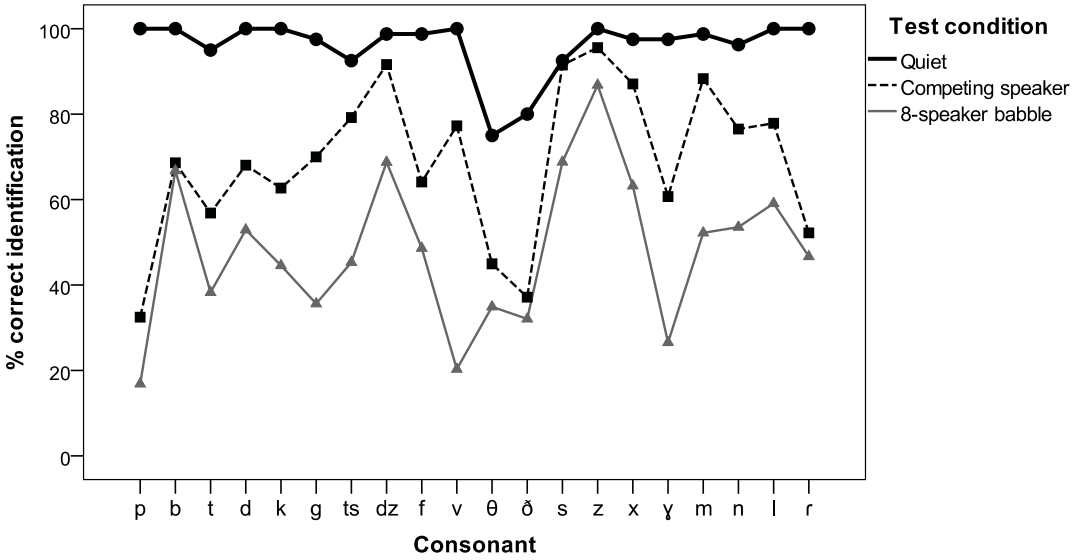


Table 1: Greek consonant confusion patterns in QUIET. Percentages of responses have been pooled over participants, consonant contexts and stress position. Responses less than 3% are not shown.

	Response																			
Stim.	p	b	t	d	k	g	ts	dz	f	v	θ	ð	s	z	x	ɣ	m	n	l	r
p	**100**																			
b		**100**																		
t			**95**	5																
d				**100**																
k					**100**															
g						**98**														
ts			7				**93**													
dz								**99**												
f									**99**											
v										**100**										
θ									25		**75**									
ð										20		**80**								
s									3				**93**							3
z														**100**						
x									3						**98**					
ɣ															3	**98**				
m																	**99**			
n																	4	**96**		
l																			**100**	
r																				**100**

consonant, $F(19,361) = 38.07$, $p<0.001$ and a significant test condition × consonant interaction [$F(46,722) = 16.16$, $p<0.001$], indicating that some Greek consonants caused more difficulty to Greek speakers than other consonants and that identification for some consonants was more affected by the presence of noise than others.

In order to examine identification scores for individual Greek consonants in more detail, confusion matrices were compiled over all speakers, stress positions and vowel contexts for each testing position and are shown in Tables 1-3. Stimuli presented are displayed in rows and listeners' responses in columns (in percentages, responses less than 3% are not shown).

Identification scores in QUIET (Table 1) ranged from 75% to 100% correct. Greek listeners identified 18/20 consonants at very high rates (>90% correct) and had difficulty identifying only two consonants, specifically the fricatives /θ/ and /ð/; both confusions were place-of-articulation confusions; dental /θ/ was confused with labiodental /f/ 25% of the times and dental /ð/ was confused with labiodental /v/ 20% of the times.

Identification scores in COMP (Table 2) ranged from 32% to 95% correct. Greek listeners identified at relatively high rates (>80% correct) /dz/, /s/, /z/, /x/ and /m/ and had most difficulties (<60% correct) with /p/ (mostly confused with /b/), /t/ (mostly confused with /d/), /θ/ (mostly confused with /f/), /ð/ (mostly confused with /v/, and /ɾ/ (mostly confused with /l/). In general, the affricates were the easiest to identify (85% correct), followed by the nasals (82% correct), the fricatives (70% correct), the liquids (65% correct) and, finally, the plosives (60% correct).

Identification scores in BABBLE (Table 3) ranged from 17% to 87% correct. Greek listeners showed moderate identification scores (>60% correct) for /b/, /dz/, /s/, /z/

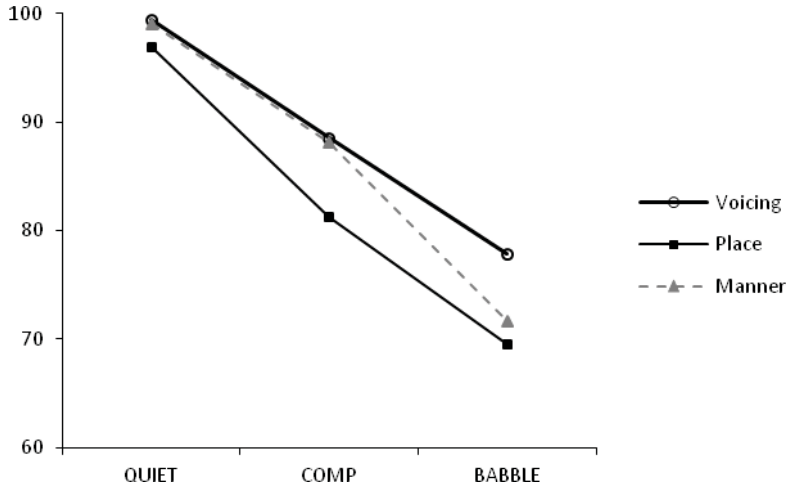
Table 2: Greek consonant confusion patterns in COMP. Percentages of responses have been pooled over participants, consonant contexts and stress position. Responses less than 3% are not shown

	Response																			
Stim.	p	b	t	d	k	g	ts	dz	f	v	θ	ð	s	z	x	ɣ	m	n	l	ɾ
p	**32**	38	3	5	4				4	3					3	3	3			
b		**68**			3	3				6					3	5	3			
t		2	**57**	28														3		
d		9		**68**								5					3	9		
k	4	3	5		**63**	11			5				3		4					
g		6	3		10	**70**				3						3				
ts			6				**79**	14												
dz				4			3	**91**												
f									**64**	6	8		7		6		7			
v		5							1	**77**		4			3	6				
θ									24	3	**45**		15	1	10					
ð				3		3				31		**37**		6	3	7				
s							3						**91**	5						
z														**95**						
x					3										**87**					
ɣ										14		3			9	**61**			4	
m																	**88**	9		
n			5									4					9	**76**		
l												3				5		3	**78**	4
ɾ		3								7		6				3	7	5	10	**52**

Table 3: Greek consonant confusion patterns in BABBLE. Percentages of responses have been pooled over participants, consonant contexts and stress position

	Response																			
Stim.	p	b	t	d	k	g	ts	dz	f	v	θ	ð	s	z	x	ɣ	m	n	l	r
p	**17**	30		4	9		3		4	3	4	3	8				3			
b	14	**66**		6														3		
t	4	13	**38**	27	6			3												
d	1	7	12	**53**		9				3		4								
k	4	3			**44**	17			5	3	5		3		4			4		
g	3	4	4	6	11	**35**	1	3		7		5	7							3
ts					3		**45**	24					15	6						
dz			3	10		3	8	**69**												
f	8	3							**48**	6	3		6	3	6		5		3	
v	4	24	3	3		3		3	3	**20**			7		3	6	10			
θ		3					3		18	3	**35**	7	13	4	8					
ð		12	4	10						15		**32**			3		9			
s									7		3		**69**	14						
z								5					3	**87**						
x					5	3			11						**63**	3				
ɣ		8		3		9			1	17		6	3	7	3	**26**				3
m		18								4		2	2				**52**	9		
n		6		3	3							3	12	3			3	**53**		
l	3	6		3									6	3				9	**59**	
r		6	6	11						6		3	3			3	3		3	**46**

Figure 4: Mean percent correct identification for voicing, place-of-articulation and manner-of-articulation in each test condition.



and /x/ and had most difficulties with /p/ (mostly confused with /b/), /t/ (mostly confused with /d/), /g/ (mostly confused with /k/), /v/ (mostly confused with /b/), /θ/ (mostly confused with /f/), /ð/ (mostly confused with /v/) and /ɣ/ (mostly confused with /v/). Again, the affricates were the easiest to identify (57% correct), followed this time by the nasals and the liquids (53% correct for both), the fricatives (47% correct), and, finally, again the plosives (42% correct).

A final analysis examined degree of difficulty as a function of the three dimensions along which consonants are characterized, namely voicing, place and manner of

articulation. Voicing had two values, voiced and voiceless. Place of articulation had four values, labial, dental, alveolar and velar. Finally, manner of articulation had five values, plosive, fricative, affricate, liquid and nasal. As shown in Figure 4, voicing proved to be a very robust feature across test conditions, followed by manner of articulation which was in turn followed by place of articulation.

4. Discussion

This paper examined the identification of Greek consonants by native listeners. The results provide a database for Greek consonant confusion patterns (including the affricates /ts/ and /dz/). The perceptual stimuli consisted of VCV tokens so that listeners could not make use of lexical information when performing the identification tasks. The VCV tokens were presented in quiet and in two types of noise, a competing speaker at an SNR of -6dB and an 8-speaker babble at an SNR of -2dB.

Greek listeners achieved excellent results across consonants in quiet (>90% correct) with the exception of the fricatives /θ/ (75% correct) and /ð/ (80% correct) (the non-sibilant fricatives are known to be difficult to identify, see e.g. Miller & Nicely 1955, Redford & Diehl 1999, Wang & Bilger 1973). Greek consonant identification deteriorated significantly in the two noise conditions. The 8-speaker babble had a greater deteriorating effect in Greek listeners' identification scores than the competing speaker despite the fact that the former had a much higher signal-to-noise ratio than latter. This result is in line with previous research (e.g. García Lecumberri & Cooke 2006, Simpson & Cooke 2005, van Dommelen & Hazan 2010) and can be interpreted as due to the fact that babble noise produces a combination of more energetic and more informational masking than a competing speaker. In the competing speaker condition, the affricates were the easiest to identify, followed by the nasals, the fricatives, the liquids and the plosives. In babble, the affricates were also the easiest to identify, followed this time by the nasals and the liquids, the fricatives, and, finally, again the plosives (see e.g. Weber & Smits 2003 for similar findings on American English listeners' difficulty in perceiving plosives in noise). Across conditions, voicing proved to be a very robust feature, followed by manner of articulation and finally place of articulation (cf. Miller & Nicely 1955, Wang & Bilger 1973, Phatak & Allen 2007, Phatak et al. 2008).

Acknowledgements

The research project is implemented within the framework of the Action «Supporting Postdoctoral Researchers» of the Operational Program "Education and Lifelong Learning" (Action's Beneficiary: General Secretariat for Research and Technology) awarded to Angelos Lengeris, and is co-financed by the European Social Fund (ESF) and the Greek State.

References

Arvaniti, Amalia. 2007. "Greek phonetics: The state of the art." *Journal of Greek Linguistics* 8: 97-208.

Benkí, José R. 2003. "Analysis of English nonsense syllable recognition in noise." *Phonetica* 60: 129-157.

Boersma, Paul, and David Weenink. 2008. "Praat: doing phonetics by computer." Accessed September 11, 2008. http://www.praat.org.

Cooke, Martin, and Odette Scharenborg. 2008. "The Interspeech 2008 consonant challenge." *Proceedings of Interspeech*. Brisbane, Australia, 1765-1768.

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: 3668-3678.

García Lecumberri, Luisa, and Martin Cooke. 2006. "Effect of masker type on native and non-native consonant perception in noise." *Journal of the Acoustical Society of America* 119: 2445-2454.

Miller, George A., and Patricia E. Nicely. 1955. "An analysis of perceptual confusions among some English consonants." *Journal of the Acoustical Society of America* 27: 338-352.

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.

Phatak, Sandeep, Andrew Lovitt, and Jont B. Allen. 2008. "Consonant confusions in white noise." *Journal of Acoustical Society of America* 124 (2): 1220-1233.

Redford, Melissa, and Randy Diehl. 1999. "The relative perceptual distinctiveness of initial and final consonants in CVC syllables." *Journal of Acoustical Society of America* 106: 1555-1565.

Simpson, Sarah, and Martin Cooke. 2005. "Consonant identification in N-talker babble is a nonmonotonic function of N." *Journal of the Acoustical Society of America* 118: 2775–2778.

Smits, Roel, Natasha Warner, James M.McQueen, and Ann Cutler. 2003. "Unfolding of phonetic information over time: a database of Dutch diphone perception." *Journal of the Acoustical Society of America* 113: 563-574.

van Dommelen, Wim A., and Valerie Hazan. 2010. "Perception of English consonants in noise by native and Norwegian listeners." *Speech Communication* 52: 968-979.

Weber, Andrea, and Roel Smits. 2003. "Consonant and vowel confusion patterns by American English listeners." *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS 2003)*, 1437-1440.

Wang, Marilyn D., and Robert C. Bilger. 1973. "Consonant confusions in noise: a study of perceptual features." *Journal of the Acoustical Society of America* 54: 1248-1266.