Mind and Second Language Acquisition

Mind and Second Language Acquisition

Experimental Approaches

Edited by

Georgios P. Georgiou

Cambridge Scholars Publishing

Mind and Second Language Acquisition: Experimental Approaches

Edited by Georgios P. Georgiou

This book first published 2022

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-8409-7
ISBN (13): 978-1-5275-8409-9

CONTENTS

INTRODUCTION

The volume *Mind and Second Language Acquisition: Experimental Approaches* consists of a collection of studies by researchers from all over the world that provide evidence about several aspects of second language (L2) acquisition using experimental protocols, or open the way for future experimental research that will allow us to better understand speakers' attitudes and motivations to learn an L2. Specifically, it concerns the perception of segmental and suprasegmental L2 features, the production of morphophonemic L2 features, the acquisition of vowel variants, the acquisition of L2 prepositions, attitudes about code-switching, and factors that motivate speakers to use and learn a heritage language. An array of different methodologies is employed which includes but is not limited to perceptual tasks, production tasks, interviews, and questionnaires. It is also of great importance that there have been studied various first or second languages some of which are underresearched. These include Cypriot Greek, English, French, Italian, Polish, Russian, Sanna (Cypriot Maronite Arabic), Scottish English, Slovak, Spanish, and Tunisian Arabic. To date, there is a need for innovative research that examines topics of L2 acquisition in an attempt to better define the mechanisms and processes that underlie the acquisition of an L2. This volume responds to this need as it offers a multidisciplinary approach to L2 acquisition and experimental methods. It also brings together researchers conducting research in the aforementioned area to discuss their research findings and exchange ideas regarding the future direction of L2 acquisition research.

Chapter 1 examines the difficulties of Cypriot Greek learners of English with respect to the perception of voicing in English plosives. Speakers of different English proficiency levels completed a forced-choice perceptual test in which they were asked to underline the word they heard from the recordings of minimal pairs with plosives. The results demonstrated that during the process of voiceless English plosives, the learners relied to a great extent on voice onset time and to some extent on the change in fundamental frequency. Furthermore, formant transitions after a plosive affected the perception of English voiceless stops by Cypriot Greek speakers. It was found that difficulties in the process of English plosives affected high proficiency learners more than lower proficiency learners.

Chapter 2 investigates the perception and production of L2 prosody by Polish bilinguals in L2 French. The perceptual test was developed with the use of prosodic morphing and addressed qualitative analyses of prosody at the interface of phonetics, phonology, syntax, and semantics. The findings yielded that cognitive processes were preventing the acquisition of L2 prosody before the highly advanced stage of L2 development. Thus, it is concluded that L2 prosody comprises the last stage of L2 acquisition.

Chapter 3 examines the assimilation of the Greek sounds [θ] and [ð] to the Russian phonological categories – these sounds are not present in the Russian system. Russian students completed a computer-based perceptual test in which they were asked to assimilate the target sounds to their L1 system. The results showed that the Greek [θ] was mostly assimilated to the Russian [s] and the Greek [ð] to the Russian [z]. These findings are not consistent with some earlier findings in the literature. The author explains that these differences might have emerged due to the learners' dialectal background, their knowledge of other foreign languages, and possible differences in their cognitive functions.

Chapter 4 investigates the production of English plural [s] in two groups of Tunisian EFL child learners differing in the English learning onset time, and the role of memorization and orthography on this acquisition. The productions were collected by means of flashcards and word lists and recorded on a PC using the Praat speech signal processing software. Children who had been learning English for 7 months pronounced better the phonetic realizations of words' final [s] compared to those who had just started to learn English. Both memorization and orthography significantly impacted the results.

Chapter 5 studies the extent to which long-term Slovak immigrant women acquire the variation of FACE and GOAT vowels compared to their typical native-speaking Scottish peers and the role of identity in their adaptation to the local language community. The data were collected through a semistructured interview, a reading passage, a word list task, and a series of tasks for comparing differences in vowel realization across speech styles. The results demonstrated that immigrants were about to acquire the monophthongal variation of /e/ and /o/, however, their productions significantly differed from the productions of the native speakers. It was also observed an association between accent acquisition and identity.

Chapter 6 examines the acquisition of L2 Italian prepositions by English monolingual and English – Spanish bilinguals. It also investigates the role of gender, experience, and task type on this acquisition. The experimental protocol required participants to complete translation, fill-in-the-blank, and multiple-choice tasks. The results showed that the L1 background of learners affected the acquisition of Italian prepositions since bilinguals outperformed monolinguals. Experience also affected the acquisition as learners with lower proficiency level presented with higher gains of accuracy in the acquisition of the prepositions. Gender and task type did not have any effect on the ability of learners in acquiring the L2 prepositions.

Chapter 7 explores the attitudes of academics in regard to the use of code-switching by Cypriot Greek speakers in EFL tertiary education classrooms and the functions it serves when it is used by them. The data were collected through questionnaires that were administered to the academics. The findings indicated that academics had mostly neutral opinions about code-switching. The author concludes that code-switching comprises an important tool in the EFL classroom since it can have pedagogical, administrative, and interactional functions. However, it should not overtake the foreign language.

Chapter 8 studies the factors which motivate Sanna speakers in Cyprus to learn their heritage language through education in adulthood. The participants of the study completed an online questionnaire. The results demonstrated that Sanna speakers associate themselves with the identity, culture, and heritage of their community. Nevertheless, they reported limited instrumental motivation regarding the use and learning of their variety. The authors propose that future experimental protocols would allow a deeper understanding of the speakers' diverse experiences.

The chapters discuss primary evidence about topics of L2 acquisition using various methodologies and speakers/learners of various languages. Their research findings are relevant but not limited to the fields of psycholinguistics, sociolinguistics, educational linguistics, and applied linguistics. Hopefully, the present volume will spark the interest of other researchers who may critically evaluate and expand the presented research or may employ the same protocols to investigate other languages or other populations.

CHAPTER 1

PERCEPTION OF L2 VOICING SYSTEM AND ITS EFFECT ON L2 SPELLING: THE CASE OF CYPRIOT GREEK SPEAKERS LEARNING ENGLISH

KATARZYNA ALEXANDER
UNIVERSITY OF NICOSIA

Abstract

Native users of Cypriot Greek (CG), a variety of Greek spoken in the Republic of Cyprus, may sporadically misspell letters representing plosives while writing in English. Users of English with CG as their L1 tend to spell words containing voiced plosives using letters that stand for voiceless counterparts of the plosives, and vice versa. Such spelling mistakes can also be found on various shop labels, signboards, menus or leaflets. Hence, it is possible to encounter spellings such as octobus instead of 'octopus', rapid not 'rabbit', delicadessen for 'delicatessen', carlic for 'garlic', or avocato for 'avocado'. This research investigates the possible link between the said spelling and possible difficulties with the perception of voicing in English plosives by speakers of CG caused by the differences between voicing patterns of English and CG. A forced-choice perceptual test was administered to a group of 90 learners of English at 5 different levels of proficiency in English. The experiment tested Greek Cypriots' perception of minimal pairs that differed in the voicing of the plosives that they contained. The analysis of the relationship between the factors affecting the perception of voicing in English plosives and the frequency of perceptual mistakes in the forced-choice test suggest that native speakers of CG do not perceive the voicing patterns of English plosives in the way native speakers of English do. It seems that native speakers of CG depend largely on VOT and to some extent on the change in fundamental frequency when processing voiceless plosives.

Also, formant transitions following a plosive play a certain role in the perception of voiceless stops by CG speakers. As far as voiced stops are concerned, the data did not reveal a clear relationship between VOT, the change in F0 or formant transitions, and the number of mistakes.

Keywords: speech perception, spelling mistakes, voice onset time, fundamental frequency, formant transitions, plosives

1. Introduction

In the Republic of Cyprus, in which the official languages are Greek (in the government-controlled areas and Turkish (in the non-government-controlled areas), English is widely used in the public domain due to the country's colonial past and to tourism, which is an important source of the country's economy. Thus, it is common to see information in shops, restaurants, businesses, hotels, etc., in English and Greek around Cyprus. The English text on various notices (such as menus, leaflets or shop labels), however, may contain spelling mistakes in words which contain graphemes that represent English plosive sounds, that is sounds such as /p/, /t/, /k/, /b/, /d/ and /g/. It is striking that the mistakes concern graphemes representing a group of sounds that are produced significantly differently in Cypriot Greek (henceforth CG, a variety of Greek used by Greek Cypriots in day-to-day conversations) and in English, and that voiced and voiceless counterparts that are normally confused.

Thus, one might pose a hypothesis that the source of confusion could be the differences in voicing between the Greek and English counterparts of the sounds. To this end, this study investigates whether and which of the known cues for voicing affect the perception of English plosives by native users of CG.

In section 2 the focus is on the description of spectral cues that are used to determine voicing in plosives, whereas section 3 gives an overview of differences in voicing patterns of CG and English. Research questions are presented in section 4, and the methods of data collection are described and discussed in section 5. Furthermore, section 6 is devoted to data analysis and discussion which include is the influence of VOT, F0 change and the first two formant transitions on perception. Finally, the data limitations are specified in section 7, whereas concluding remarks are presented in section 8.

2. Acoustic cues for voicing

There is a plethora of phonetic experiments which indicate that several acoustic cues can influence listeners' decisions concerning whether the stop consonant they hear is voiced or voiceless. An account of some of these cues, such as variation in fundamental frequency, the length of voice onset time and the transition of the first two formants following the release of the plosive, is given in this section.

Lisker and Abramson (1964), who analysed acoustically initial plosives across languages, found that voice onset time (VOT) values, i.e. the time between the offset of a burst of a plosive and the beginning of voicing, can be utilised for categorising voicing manners (such as voiced or voiceless) in plosives. The study also showed that languages can have from two to four different voicing categories for each place of articulation; that is, voiced, voiced aspirated, voiceless unaspirated and voiceless aspirated. Additionally, the data suggest that the length of VOT also depends on the place of articulation of the given plosive. The relation of the length of VOT and the place of articulation was also in agreement with previous findings by Liberman, Delattre and Cooper (1958). The findings by Lisker and Abramson (1964) as to the status of VOT as a spectral cue of voicing were consistent with studies which employed tests using synthetic speech by Liberman, Delattre and Cooper (1958), as well as Lisker and Abramson (1970).

Stevens and Klatt (1974) indicated the importance of the timing of formant transition after the release of a plosive as well as a trading relationship between VOT and formant transition in the perception of voicing in plosives. The data from a set of experiments by Stevens and Klatt (1974, 657) on the perception of synthesised syllables by English speakers indicated that for a plosive to be perceived as voiceless the transition of the first two formants needs to take place before the voicing starts. Also, there has to be minimum 20ms of VOT. In case the transition continues after the voicing onset, listeners tend to rely on different cues, that is, either on the VOT or the speed of formant transition. Thus, having a long VOT increases the chances of a plosive being identified as voiceless by some listeners; however, voicing onset with a rapid change in the frequency of the first two formants may signify voicing for other listeners (Stevens and Klatt 1974, 657). Nonetheless, data coming from research on the perception of synthesised syllables by native English speakers carried out by Lisker (1975) suggest that F1 transition is not essential for plosives to be perceived as voiced and that it is not certain whether it is necessary for initial plosives not to have F1 transition to be heard as voiceless. He also draws attention to the tendency among many languages (e.g. Spanish) not to use F1 transition as a cue for voicing, which would make it very unusual for F1 to be used as a cue in English (Lisker 1975, 1550).

Research carried out by Ohde (1984) demonstrated that fundamental frequency values differ for voiced, voiceless unaspirated and voiceless aspirated plosives in American English. Specifically, voiceless aspirated and voiceless unaspirated plosives have a much higher fundamental frequency than voiced plosives at all periods following offset. The difference between voiceless aspirated and unaspirated plosives is that the fundamental frequency at the first period is substantially higher for unaspirated plosives than for the aspirated ones, where at the following periods the values for aspirated and unaspirated plosives were very similar. As far as the rate of change of frequency in the first two periods is concerned, the research by Ohde (1984) revealed that the change depends on the place of articulation, voicing and the adjacent vowel. It was discovered that voiceless unaspirated stops undergo the greatest changes of all the stops, whereas aspirated stops have generally similar F0 values to voiced stops for labial and alveolar places of articulation, but higher F0 values than voiced for velars. One should note that although English voiceless unaspirated plosives and voiced plosives are very similar with regard to VOT values, they differ substantially with respect to the fundamental frequency (Ohde 1984, 226 – 227).

3. Voicing patterns in Cypriot Greek and English

According to several sources (Armosti 2010; Arvaniti 2010; Newton 1972), there are only voiceless plosives in the phonemic inventory of CG. CG voiceless plosives, however, have a few allophonic variations. If preceded by a homorganic nasal, /z/, vowel or sonant, the stops surface as voiced. While Newton (1972) classifies voiceless geminate plosives as allophones, later descriptions of CG plosive inventories (Armosti 2010; Arvaniti 2010) state that this group of sounds should be treated as phonemes, especially that they may occur in the same positions as singleton voiceless stops, and there are minimal pairs in CG in which the only different sounds are singleton and geminate stops. Nevertheless, the discussion of what is a phoneme or an allophone has no significance for this study. What is noteworthy is what types of plosives can be found in certain contexts in CG.

Table 1. Voicing patterns of CG and English: a summary

	Cypriot Greek	English
Voiced	**Pre-nasalised** [^{m}b, ^{n}d, ^{ŋ}g]	**Short VOT** (around 10ms) **Or negative VOT** (Ladefoged 2001)
Voiceless	**Short VOT** /p/ - 14ms, /t/ - 16ms, /k/ - 31ms (Arvaniti 2001)	**Long VOT** (/p/ - 50 - 60ms) (Ladefoged 2001)
Voiceless geminate	**Long VOT** /p:^{h}/ - 50ms, /t:^{h}/ - 60ms, /k:^{h}/ - 63ms + **longer closure duration** (Arvaniti 2001)	

Several measurements of VOT values in CG have been made so far. For instance, according to Arvaniti's research (2001, 41, Figure 2), the mean VOT durations of singleton plosives at normal speech rate are around /p/ - 14ms, /t/ - 16ms and /k/ - 31ms. There is no change in VOT value for /p/ and /t/ at a fast speech rate, however, in the case of /k/ the duration falls to about 18ms. The speech rate seems to have a great influence on the VOT values of geminates. In line with Arvaniti's study (2001, 41), the mean VOT durations are about 50ms for /p:^{h}/, 60ms for /t:^{h}/ and 63ms for /k:^{h}/ at normal rate, whereas at fast rate the mean durations fall to nearly 40ms for /p:^{h}/ and around 45ms for /t:^{h}/ and /k:^{h}/. A similar study by Arvaniti and Tserdanelis (2001, 4, Figure 2) reveals the VOT values may also change depending on whether a geminate plosive is in a stressed or unstressed syllable, namely in stressed syllables, average VOT values seem to be about 75ms for /t:^{h}/ and /k:^{h}/, and about 60ms for /p:^{h}/. To sum up, the VOT of singleton voiceless stops has been asserted to range from about 14ms up to 31ms, while the VOT durations for geminate plosives start from 40ms and reach 75ms.

As claimed by Ladefoged (2001, 127), the VOT duration of English stressed /p/ in word-initial position ranges between 50 and 60 ms. As far as /b/ is concerned, it may have up to about 10ms of VOT, however, the value can also be negative. As regards /p/ following /s/, usually described as unaspirated, it is characterised by VOT value similar to that of /b/ (Ladefoged 2001, 127). Taking into consideration the findings by Newton (1972), Arvaniti (2001), Arvaniti and Tserdanelis (2001) and data provided by Ladefoged (2001), it can be maintained that as to VOT values, CG voiceless plosives seem to be more like English voiced and unaspirated plosives, than to aspirated voiceless plosives. CG geminate plosives in stressed syllables (syllable initial position) range between 60 and 75ms, which in this respect makes them similar to English stressed voiced stops (between 50- 60ms). CG voiced stops, unlike their English counterparts, are obligatorily pre-nasalised.

4. Research questions

Taking into consideration the observations discussed in section 1 regarding the potential perception-related spelling difficulties characteristic of learners of English whose L1 is Cypriot Greek, the differences in plosive inventories, and the various cues possibly used by listeners to determine the voicing of a given sound, the following question was put forward:

1. Could the differences in the acoustic properties in plosive inventories cause speakers of Cypriot Greek as L1 difficulties with their 'perceptual fine-tuning' of English plosives? And if so, two related questions were asked:
   1a. In what word contexts do listeners whose L1 is Cypriot Greek have difficulties with precise identification of voicing patterns in English?
   1b. Which acoustic cues have an influence on the perception of voicing in English plosives by listeners whose L1 is Cypriot Greek?

5. Method of data collection

The experiment tested Cypriot Greek speakers' perception of minimal pairs that differed in the voicing of the plosives that they contained. The plosives appeared in various positions in words, which were selected in line with different ways of realisation suggested by Gimson (1989); the positions of tested plosives are specified in Tables 2 and 3 in section 6. Each of the words in each minimal pair was embedded in a sentence (see Appendix 1). The sentences were written in a way that allowed both words in each minimal pair to be used in a relatively contextualised lexical chunk. A recording of an adult male British-English native speaker (standard southern variety) reading the sentences was made. Then, each of the sentences was copied and pasted using PRAAT in a way that enabled the same sentence to be played twice. Afterwards, all the recordings were converted into WAV format and a list of sentences (each repeated twice) was created.

The list of sentences was played to several different language-level groups of adult English learners. At the time of listening, each of the subjects could see the list of previously recorded sentences on a worksheet that presented each sentence including both minimal-pair words printed in bold. The task of the subjects was to underline the word in each minimal pair that they heard.

The perceptual task was carried out in quiet classrooms. The recordings were played through two loudspeakers.

6. The results

The total number of mistakes was 21.6% for all levels i.e. this was the percentage of students that chose the wrong minimal pair word for a given recorded sentence. Such a result appears to indicate that there is a difference in the way native speakers of CG and of English perceive voiced and voiceless plosives in English words produced by a native user of English. Furthermore, it was found that the number of mistakes was lower for voiceless targets (17.6%) than for the voiced ones (25.66%) (see Tables 2 and 3). There was noticeable variation within the ‘voiced’ and ‘voiceless’ categories as well. That is, with regard to voiced plosives (Table 3), the lowest percentage of errors was for velars (22.98%), and the highest percentage of errors was for bilabials (25.9%). Whereas, with reference to voiceless stops (Table 2), the highest percentage of mistakes occurred in velars (19.9%), the lowest percentage of mistakes was for alveolars (14.3%) and bilabials had intermediate values (18.6%).

Table 2. Overall percentages of mistakes for voiceless plosives.

	Accented aspirated		Weakly accented, Relatively unaspirated		Syllable Final	Followed by nasal consonant	Followed by lateral consonant	**Total for of mistakes each class**
	Followed by vowel	Followed by a consonant	Following a vowel	Following a nasal				
p	pet	pride	Mopping	simple	rip	Lapmate	nipple	
%	4.4	5.5	7.7	33.3	15.5	56.6	7.7	**18.6**
t	tied	-	Latter	renter	bet	mitten	petal	
%	14.4	-	10	11.1	33.3	5.5	6.6	**14.3**
k	cards	class	Locker	anchor	back	Pickman	ankle	
%	4.4	2.2	6.66	36.6	40	23.3	26.6	**19.9**
TOTAL (for all voiceless plosives)								**17.6**

Table 3. Overall percentages of mistakes for voiced plosives.

	Initial, Partially devoiced		Intervocalic Voiced		Final, voiceless	Followed by a nasal consonant	Followed by a lateral consonant	**Total for of mistakes each class**
	Followed by a vowel	Followed by a consonant	Following a vowel	Following a consonant				
	bet	bride	mobbing	symbol	rib	Labmate	nibble	
b	14.4	6.6	35.5	16.66	51.1	48.8	8.88	**25.9**
	dyed	-	Ladder	render	bed	midden	pedal	
d	12.2	-	25.5	5.5	42.2	27.7	55.5	**28.1**
	guards	glass	Logger	anger	bag	Pigman	angle	
g	1.1	11.1	16.66	15.5	22.2	65.5	28.8	**22.98**
TOTAL (for all voiced plosives)								**25.66**

The data in Table 4 suggest that there is a relationship between student language level and perceptual development. That is, lower-level students appeared to have more perception-related mistakes than higher-level students. However, there was one exception, for the overall score for B2[1] level students, which might have been slightly skewed by the small sample (only 9 students).

Table 4. Percentages of mistakes and numbers of subjects according to the level of English. [2]

Level	A2	B1	B2	C1	C2	**TOTAL for all the students**
Number of subjects	9	32	9	24	16	**90**
%	27.2	24.5	19.4	20.3	14.2	**21.6**

Study data also suggest that there was a tendency for sounds in certain word contexts to have caused a relatively high level of mistakes. This was true for voiced and voiceless plosives in word-final position, compounds and words in which a plosive followed a nasal. It may be the case that the mistakes in the words with a plosive following a nasal could have been made not only because of perception. To be more precise, in Standard Modern Greek, which is the official language in Cyprus, in any word-medial sequence in which a plosive follows a nasal, the plosive is realised as voiced (Newton 1972, 22-23). Based on anecdotal observations, I tentatively hold that sequences that are spelled with μπ, ντ and γκ often appear to be treated as equivalents of *nk*, *nt*, *mp* or even *nch* in the Latin alphabet by Greek and Greek Cypriot learners of English as L2. It is worth considering then, that some of the subjects might have attached more importance to the spelling of the word while taking part in the experiment, or even perceive a sound as voiced depending on the cue which was a nasal element before a plosive, as this would be the case in their native language.

---

[1] In this paper reference will be made to Common European Framework of Reference (CEFR). B2 level is an upper-intermediate level

[2] The levels of the courses are thought to be approximate to the ones specified by CEFR.

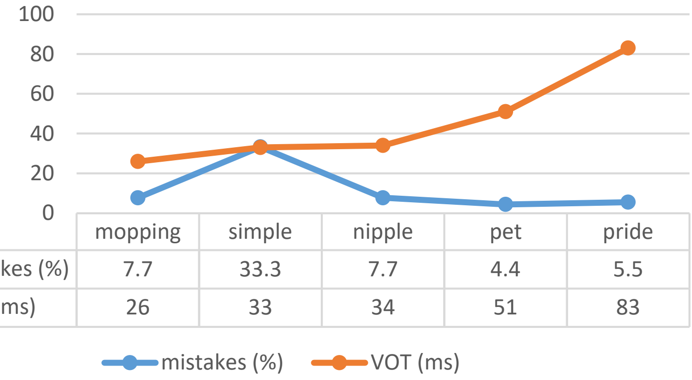
6.1. Data analysis I: Perception of voiced and voiceless plosives with different VOT values

6.1.1. Voiceless plosives

Table 1 shows percentages of mistakes for all voiceless plosives tested in the study. Some of the words did not undergo further analysis as a result of problems with measuring the VOT in the stop consonants that they contained. To be more precise, in this study VOT is defined as the time between the offset of the plosive and the beginning of vibration of vocal cords (for voiceless plosives) or the time between the offset of burst and the first visible peak of a periodic wave (for voiced plosives). Thus it could not be measured for the sounds that were in word-final position and were are not followed by any other voiced sound. Therefore, words like *rip*, *bet*, and *back* were not included in my analyses of the relationship between the length of VOT and the number of minimal-pair mistakes. In addition, it was not possible to detect VOT in the word *Lapmate*, so it was not incorporated in further analyses considering the length of VOT.

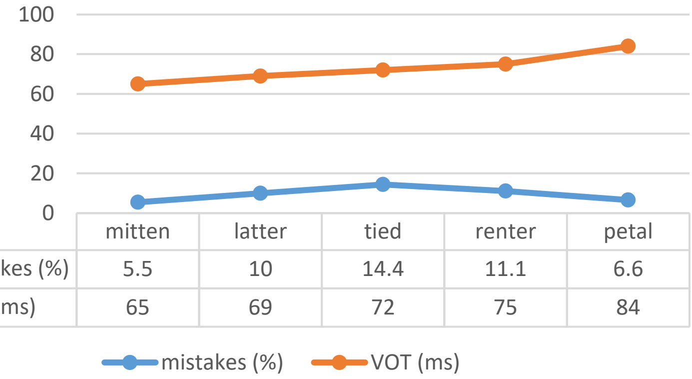
Figures 1, 2 and 3 below illustrate the relation between the English native speaker's VOT values for plosives in given words and the percentages of subjects' mistakes made in the minimal pair test. Results for /p/ show that for VOT ranging from about 26ms to around 34ms, the number of subject mistakes was 7.7% for *mopping* and *nipple*, and 33.3 % for *simple* which might have had a greater number of mistakes due to the *nasal* + *plosive* sequence. Yet, from 51ms up to 83ms of VOT the number of mistakes decreases as the VOT goes up. For items with /t/, as VOT increases from 65ms to 75ms, the number of mistakes slightly rises but starts falling for VOT values between 75 and 84ms. From the data on alveolars and bilabials, it is possible to notice a tendency where, if VOT is longer than 51ms for /p/ and 84ms for /t/, the percentage of mistakes is not very high. However, the data for /k/ seem to be more varied. The words *anchor* and *ankle* (*nasal* + *plosive* sequence) and *Pickman* (compound) caused a relatively higher percentage of mistakes than the other currently analysed words with /k/. However, even if the three items are not taken into consideration in the analysis, the same tendency for /p/ and /t/ did not seem to apply. The reason for this situation may be that there may have been some additional factors affecting the results (see sections 6.2. and 6.3.).

Figure 1. The relationship between the percentage of mistakes and the length of VOT in items with /p/



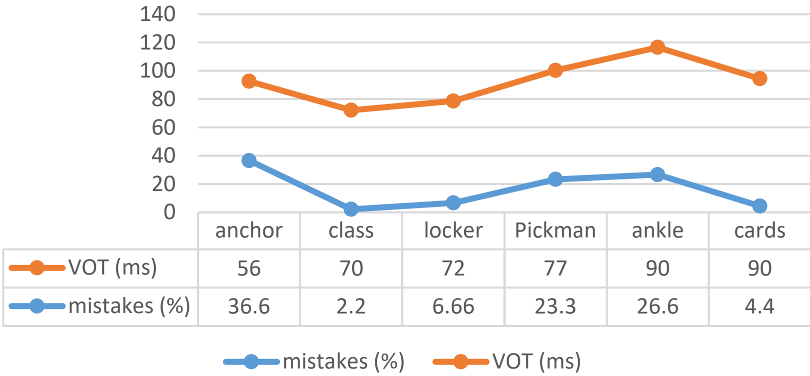
	mopping	simple	nipple	pet	pride
mistakes (%)	7.7	33.3	7.7	4.4	5.5
VOT (ms)	26	33	34	51	83

Figure 2. The relationship between the percentage of mistakes and the lenght of VOT for the items with /t/



	mitten	latter	tied	renter	petal
mistakes (%)	5.5	10	14.4	11.1	6.6
VOT (ms)	65	69	72	75	84

Figure 3. The relationship between the percentage of mistakes and the length of VOT for the items with /k/



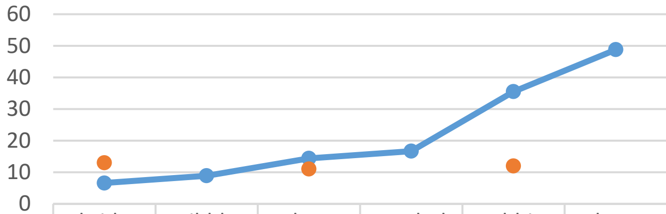
	anchor	class	locker	Pickman	ankle	cards
VOT (ms)	56	70	72	77	90	90
mistakes (%)	36.6	2.2	6.66	23.3	26.6	4.4

6.1.2. Voiced plosives

Figures 4, 5 and 6 below present the link between the percentage of mistakes and VOT in the voiced plosives examined. To investigate this link, a different definition of VOT was applied, namely the distance from the offset of the plosive to the first visible peak of a periodic wave on a waveform[3]. This measurement was taken irrespective of whether any of the items was prevoiced[4]. In several of the items examined in this study, the offset of the plosives could not be seen clearly in the waveform (i.e. *Labmate, midden, nibble, symbol, Pigman*), therefore the VOT was not specified for these words, and thus, they appear in Figures 4, 5 and 6 without the value for VOT. As measuring VOT for voiced stops in word-final position was not possible, the words *rib*, *bed* and *bag* were not included in the following analysis. Information on the percentage of mistakes for these words can be seen in Table 3.

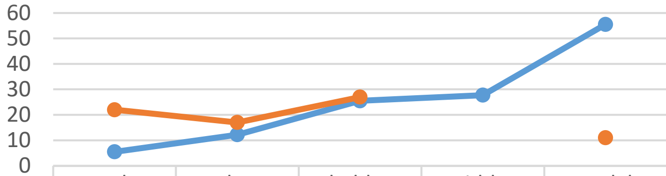
[3] Such procedure was earlier employed by Ohde (1984).
[4] The prevoiced items were as follows: *anger, dye, midden, mobbing, nibble* and *pedal.*

Figure 4. The relationship between the percentage of mistakes and the length of VOT for items with /b/



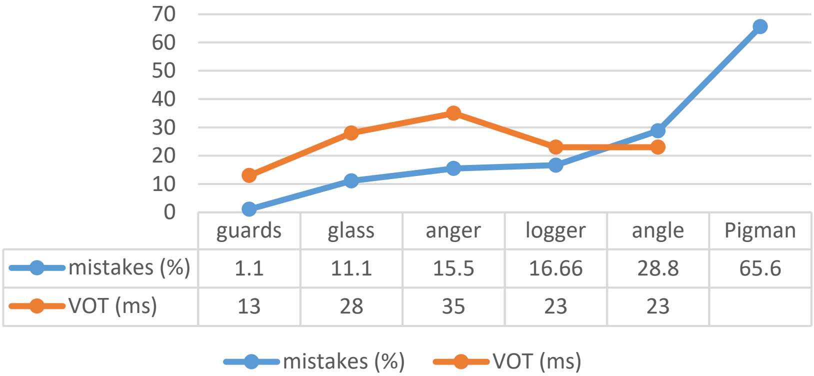
	bride	nibble	bet	symbol	mobbing	Labmate
mistakes (%)	6.6	8.88	14.4	16.66	35.5	48.8
VOT (ms)	13		11		12	

Figure 5. The difference between the percentage of mistakes and the length of VOT for items with /d/



	render	dye	ladder	midden	pedal
mistakes (%)	5.5	12.2	25.5	27.7	55.5
VOT (ms)	22	17	27		11

Figure 6. The difference between the percentage of mistakes and the length of VOT for items with /g/



	guards	glass	anger	logger	angle	Pigman
mistakes (%)	1.1	11.1	15.5	16.66	28.8	65.6
VOT (ms)	13	28	35	23	23	

The comparison of VOT values for voiced plosives with corresponding percentages of mistakes did not reveal any clear correlation. High percentages of mistakes were found for short and long VOTs and even if prevoicing of some of the items is taken into consideration, no clear pattern seemed to emerge from the data. This study's data suggested that the perception of voicing in the case of voiced stops did not seem to be dependent on VOT only, if at all.

The length of VOT did not seem to be the only factor in difficulties in the accurate perception of voiceless plosives and did not appear to have much influence, if any, on the correct perception of voiced plosives. Therefore, the influence of changes in fundamental frequency and formant transitions were analysed in sections 6.2. and 6.3. respectively.

6.2. Data analysis II: Perception of voiced and voiceless plosives with different fundamental frequency values

As there appeared to be more reasons for the variation in the percentages of mistakes in voiced and voiceless plosives than the length of VOT, an examination of the influence of the fundamental frequency on the perception of the target sounds was undertaken. The fundamental frequencies of the first and second glottal period were measured for every word, as the most substantial differences between voiced, voiceless aspirated and voiceless unaspirated plosives seem to occur in the change between the frequency of the two cycles (Ohde 1984, 226-227). The length of each cycle was measured by placing cursors on the peaks, as well as at zero points of each glottal cycle. The frequency then was calculated as a reciprocal of the length of the glottal period[5]. A change in frequency between the first and the second glottal period for each word was then calculated and compared with the number of mistakes in a minimal pair test and VOT values. In order to facilitate data analysis, instead of giving the actual values of F0 change, corresponding names of categories from Ohde's data were used.[6]

6.2.1. Voiceless plosives

The results presented in Table 5. show that the F0 change might also have had an influence on the perception of voicing in plosives. Some trends could be observed, when the items in Table 5. were grouped into three categories according to the percentage of mistakes (Category I – from 2.2% to 6.6%, category II – from 7.7% to 14.4% and Category III – from 15.5% to 56.6%). The items in Category I seem to have very low numbers of mistakes due to either long VOT and/or the F0 change characteristic of voiceless aspirated or unaspirated stops. There are two exceptions (*cards* and *petal)* which have F0 change like voiced plosives. It seems, however, that their very long VOT is enough for most of the speakers to consider the items voiceless. In the second class it is possible to find items with plosives of three types: 1. not very long VOT (26-34ms) and F0 change characteristic of voiceless aspirated and unaspirated plosives, 2. long VOT and fundamental frequency change as in voiceless unaspirated plosives and 3. long VOT and the change in F0 like for voiced stops. Category III contains three out of four words with the nasal + plosive sequence, compounds and items in which the VOT

---

[5] It was decided in this analysis not to extract F0 by means of PRAAT as the program measures F0 only for a specified point in time as opposed to a period of time. Therefore, it is only possible to measure the F0 at a certain point of a glottal period, while the F0 value is likely to change within the same period, which would pose the problem of which value should be chosen for the analysis. Hence, because the results were meant to be analysed on the basis of the results achieved by Ohde, choosing Ohde's procedure seemed to be the most advisable.

[6] For example, the F0 change for *pet* was 17Hz and in Ohde's (1984) study it was found that that value within the class of bilabials was characteristic of voiceless aspirated plosives, hence instead of the value the category 'voiceless aspirated' was assigned.

and F0 change could not be measured due to the offset of burst not being clearly visible on the spectrum and waveform, or the lack of a voiced element following the stop consonant. These words caused relatively large numbers of mistakes.

Table 5. The relationship between the percentage of mistakes, VOT and the degree of change in F0 for voiceless plosives.

Category	Word	Mistakes (%)	VOT (ms)	Voicing category according to F0 change
I	Class	2.2	86	Unaspirated
	Cards	4.4	90	Voiced
	Pet	4.4	51	Aspirated
	Pride	5.5	83	Aspirated
	Mitten	5.5	64	Unaspirated
	Petal	6.6	84	Voiced
	Locker	6.6	71	Unaspirated
II	Mopping	7.7	26	Aspirated
	Nipple	7.7	34	Unaspirated
	Latter	10	69	Unaspirated
	Renter	11.1	75	Voiced
	Tied	14.4	72	Unaspirated
III	Rip	15.5	-	-
	Pickman	23.3	-	-
	Ankle	26.6	90	Aspirated
	Simple	33.3	37	Aspirated
	Bet	33.3	-	-
	Anchor	36.6	56	Aspirated
	Back	40	-	-
	Lapmate	56.6	-	Aspirated

6.2.2. Voiced plosives

Table 6. The relationship between the percentage of mistakes, VOT and the degree of change in F0 for voiced plosives.

Word	Mistakes (%)	VOT (ms)	Voicing category according to F0 change
Guards	1.1	13	Voiced
Render	5.5	22	Unaspirated
Bride	6.6	13	Unaspirated
Nibble	8.8	-	Voiced
Glass	11.1	28	No category; closest to unaspirated
Dye	12.2	17	No category; closest to unaspirated
Bet	14.4	11	Voiced
Anger	15.5	35	Voiced
Logger	16.6	23	Voiced
Symbol	16.6	-	-
Ladder	25.5	27	Voiced
Midden	27.7	-	
Angle	28.8	23	Voiced
Mobbing	35.5	12	No category; closest to unaspirated
Labmate	48.8	-	No category; closest to voiced
Pedal	55.5	11	voiced
Pigman	65.6	-	-

A summary of relationships between mistake percentages, VOT and the change in F0 for voiced plosives (Table 6) indicates that there are three options of how F0 changed in the items in this study. That is, the variations corresponded to what in Ohde's (1984) study was labelled as voiced plosives, voiceless unaspirated plosives or the changes that did not match any category. In cases where the rate of change did not match any category, the values were close to those of voiceless unaspirated plosives. For some of the items, (*symbol, midden, Pigman)* measurement of F0 change was not possible. Nevertheless, it is difficult to observe any clear indications as to whether any of the degrees of change in F0 and their corresponding voicing categories (according to Ohde 1984) have any influence on the number of mistakes. That is, items that caused more than around 14% of the mistakes, generally match the category of voiced plosives, excluding *symbol, midden* and *Pigman,* which could not have had their F0 changes measured, and *mobbing* (closest to unaspirated*)* and *Labmate* (close to voiced). On the other hand, the items that had around 12% and below would fit all the categories mentioned i.e. voiced, aspirated and unaspirated. What is more, excluding any problematic items such as those with *nasal* + *plosive* sequence or compounds from the analysis would not result in receiving any clear pattern of the influence of F0 change in voiced plosives on their perception by native speakers of CG.

6.3. Data analysis III: The influence of formant transition and VOT on the perception of voicing.

The following analysis was aimed at detecting possible influences of the timing of the first two formant transitions on the perception of voicing in English plosives by native speakers of CG. For all plosives in this study, the length of formant transition was measured from the offset of the plosive (identified on the waveform displayed in PRAAT) up to the point where a substantial transition in the formants was completed (on the spectrogram in PRAAT). In some of the words, the offset of the plosive was not possible to detect, and therefore VOT and formant transitions were not measured.

As can be seen in Figure 7, there are no items that were known to have a VOT longer than 20ms and completed formant transition by the beginning of voicing at the same time, which as Stevens and Klatt (1974) claimed would make them more likely to be perceived as voiceless by native speakers of English. In nearly all items, formant transitions took longer than the distance from the offset to the first visible peak of a periodic wave. Neither the length of formant transitions from the burst nor the length of formant transitions after the beginning of voicing (see Figure 8) seemed to have any correlation with the percentage of mistakes. Those data correspond with the findings of Stevenson and Klatt (1974) according to which listeners tend to rely on either formant transitions or VOT while judging if a plosive is voiced or not, unless the VOT is longer than 20ms and the formant transitions are completed by that time. Also, it is important to remember that that data are relevant to native speakers of English, and might differ for speakers of CG.

Figure 7. Percentages of mistakes, VOT and the timing of the formant transitions for voiced plosives

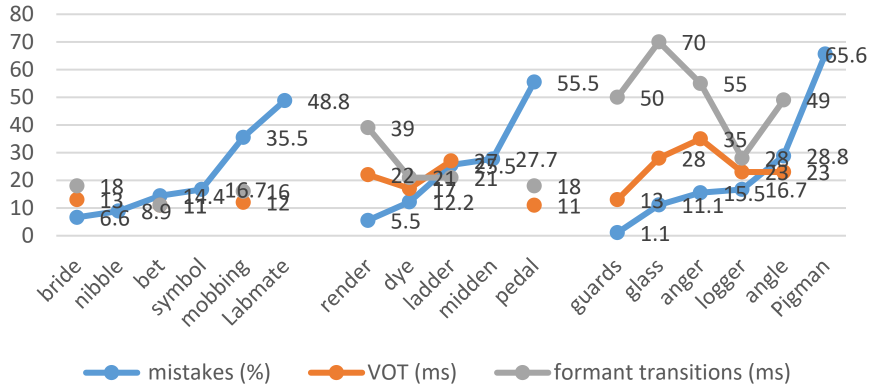


Figure 8. Percentages of mistakes, VOT and the formant transition timing after the beginning of voicing in voiced plosives

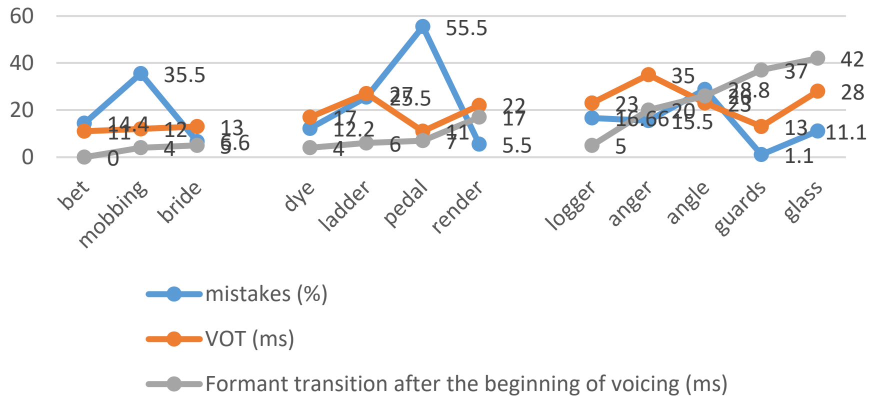


Table 7. The relationship between the percentage of mistakes, VOT and the time of formant transitions of voiceless plosives.

	Mistakes (%)	VOT (ms)	Formant transition (ms)
Mopping	7.7	26	39
Nipple	7.7	34	b[7]
Simple	33.3	33	96
Pet	4.4	51	b
Pride	5.5	83	b
Lapmate	56.6	53	
Rip	15.5		
Mitten	5.5	65	b
Latter	10	69	76
Tied	14.4	72	102
Renter	11.1	75	b
Petal	6.6	84	b
Bet	33.3		
Anchor	36.6	56	71
Class	2.2	70	100
Locker	6.66	72	b
Pickman	23.3	77	b
Ankle	26.6	90	112
Cards	4.4	90	129

A large number of voiceless plosives were found to have their VOT longer than 20ms and formant transitions completed by that time (see Table 7). However, there were a few items in which transitions of formants took longer to complete than it took voicing to begin. The relationship between the number of mistakes, length of formant transition and VOT are presented in Figure 9. At first glance, the data seem to suggest that the longer the transition of formants, the greater the number of mistakes. The picture seems to be spoilt by the items *ankle* and *anchor*, which do not fit this pattern, despite having shorter formant transitions more mistakes were made with these words than with other words containing /k/. If the reason for this were the *nasal + plosive* sequence in both of the items, the same should have applied to *simple*. If *simple* were excluded from the analysis, there would be only one item left in the bilabial category following the pattern,

[7] b – transition of formants completed before the beginning of voicing.