Universitat de les Illes Balears

Facultat de Filosofia i Lletres

Memòria del Treball de Fi de Grau

The Impact of Masks in the Intelligibility of English as a Foreign Language

Paula Bauzá Mas

Grau d'Estudis Anglesos

Any acadèmic 2020-21

DNI de l'alumne: 43220797W

Treball tutelat per Lucrecia Rallo Fabra.
Departament de Filologia Espanyola, Moderna i Clàssica.

	Autor		Tutor	
S'autoritza la Universitat a incloure aquest treball en el Repositori Institucional per a la seva consulta en accés obert i difusió en línia, amb finalitats exclusivament acadèmiques i d'investigació	Sí	No	Sí	No
	☒	☐	◪	

Paraules clau del treball:
Speech perception, auditory-visual perception, articulation, consonants, minimal pairs, masks

Index

Abstract

The Covid-19 Pandemic has forced society to change several habits and adopt health measures that have affected our daily lives. In the educational system, the obligation of masks in the class environment has affected teaching methodologies. This paper aims at analysing the effects of wearing a mask in the field of English as a Foreign Language. An experiment to a reduced group of secondary school students has been conducted to prove the effects masks have on auditory and articulatory perception. With the participation of ten students, the paper demonstrates that they experienced slight difficulties in the perception of minimal pairs containing the voiced bilabial plosive /b/ and the voiced labiodental fricative /v/ when they were uttered in a mask condition. Furthermore, the paper provides an overview of the previous literature that deals with speech perception and concludes with an approach of the implications the experiment has in the communicative aspect of secondary education.

---

1 Introduction

Listening skills are crucial to achieve effective communication in a foreign language. For this reason, speech perception has been studied for many years, and a variety of theories have emerged trying to shed light to the mysteries of how humans understand language. One of the main concerns that linguists have tried to decode deals with the relationship between auditory and visual categories in the field of perception. “In normal speech communication, visual and auditory cues are consistent, and listeners use both in making phoneme judgments.” (Diehl, Lotto, and Holt 2004, 168). Similarly, as the scholars Erdener and Burham argue, “speech perception is an auditory–visual event; when the auditory signal is degraded, visual speech information enhances the percept.” (2013, 121). The Covid-19 pandemic has made the use of masks a necessary item for our daily lives. Thus, it is evident that the development of this skill has recently been challenged by the introduction of masks. Masks cover our articulatory system entirely and this might alter the way in which listeners perceive different sounds and lead to misunderstandings. Recent investigations around this issue will be later developed. Important factors to take into account in order to study this auditory-visual relation are an overview of the main theories dealing with speech perception and the articulatory characteristics of the different sounds in the English language.

1.1 Speech Perception

Following the work by Diehl in collaboration with other researchers, they offer a brief explanation about three main theories of speech perception. One of them is the Motor Theory of Speech Perception (MT), an early theory that suggests the objects of speech perception as being articulatory events rather than auditory events. Such events are "neuromotor commands to the articulators (e.g., tongue, lips, and vocal folds)—also referred to as intended gestures" (Diehl, Lotto, and Holt 2004, 150). Similarly, the Direct Realist Theory of speech perception (DRT), agrees with the articulatory notion of perception but, in contrast to the MT, asserts that articulatory objects are "phonetically structured, vocal tract movements, or gestures, and not events that are causally antecedent to these movements, such as neuromotor commands or intended gestures" and denies the role of human-specific mechanisms (2004, 152). These scholars finally reflect on the Categorical Perception, which has defining properties that imply "speech perception discriminability is closely related to the presence or absence of functional (i.e., phonemic) differences between sounds" (2004, 155-6). Nevertheless, even though we have these theories, there is still uncertainty regarding this field of study and there are a huge number of academic articles discussing around the field.

In terms of previous literature focusing more directly into foreign language learning, Erdener reports the findings in studies carried out among Spanish speakers learning English. As the author reflects, "English differs significantly from Spanish in its phonological repertoire and most native speakers of Spanish confuse English phonemes that are not present in Spanish in both perception and production tasks" (Erdener 2012, 2). This would make evident a reduction in terms of perception errors when auditory-visual input is presented, rather than only auditory (2012, 3). Generally, in all the experiments that have been done in the context of foreign languages, listeners showed an increase in the visual speech influence (2012, 2). This makes evident the relationship or the interaction between visual and auditory aspects of communication. The following paper will offer a new approach on the topic and will try to explain what happens when the visual input is covered by masks.

Apart from the mentioned theories, there are some experiments that are relevant for this paper and connect with the theoretical framework of speech perception. One of these experiments is the "McGurk effect".

1.2 "McGurk effect"

The “McGurk effect” was a project that emerged around forty years ago and, recently, John McDonald has published a review in which he explains the background of the experiment that he developed together with Harry McGurk in 1976. This event had a great impact in the field, and it has been quoted in many following studies. The purpose of the project had the focus on young infants and their ability to coordinate visual and auditory activity in perception during the first year of life (McDonald 2018, 8). Erdener and Burnham explain that “in a typical demonstration of the McGurk effect, when the auditory syllable /ba/ is dubbed onto the lip movements for /ga/, the resultant percept among most native English speakers is ‘‘da’’ or ‘‘tha’’”(Erdener and Burham 2013, 121). This explains that the listener was influenced by the auditory input and also the articulatory input. Since they did not match, the resulting sound was different too. The experiment has enabled researchers to think of speech perception as more related with the vocal tract activity and the illusion provoked by the McGurk effect “shows that facial speech information is not simply an adjunct of auditory speech but is an intrinsic component of normal speech perception” (McDonald 2018, 16). The multimodal nature of speech perception has also moved the focus of study in visible speech, making it the compensatory information for hearing-impaired (2018, 13). The multisensory aspect of speech perception, to a great extent discovered throughout the McGurk effect, works as a kind of basis for the following study. Taking into account that wearing masks erases the possibility to visualize the vocal gestures, speech perception might be altered, and participants might not be allowed to assert the correct consonant sound.

1.3 Articulatory characteristics of consonants

According to Ladefoged (2014), the movements produced by the vocal organs have been traditionally related to the description of speech sound. The main structures are the lungs, the respiratory system and the vocal organs. Similarly, Coelho and others follow the same explanation by adding that speech production “requires proper interaction between different structures of the vocal tract, such as maxillomandibular osseous bases, dental arches, teeth, and hard palate; and of soft tissue: soft palate, tongue, lips, cheeks, and organic spaces . It is a complex task, which needs to be performed in coordinated, organized and planned fashion in order to guarantee full comprehension of what had been said by the interlocutor” (Coelho, Vieira, Bianchini 2019, 2). Figure 1 shows the different elements of the vocal tract.

Focusing on consonants, they are classified depending on manner of articulation and place of articulation. On the one hand, referring to the place of articulation, every time we

produce any consonant, an obstruction in the airstream takes place in the vocal tract and the place of such blocking allows classification of consonants (Ladefoged 2014). In the case of the present study, that deals with consonants /b/ and /v/, the places of articulation concerned are the lips and the teeth, bilabial for /b/ and labiodental for /v/. It is important to recognise this distinction because it produces certain movements on the speaker's mouth that would be hidden when the mask is present, hence, it is possible that this would have certain repercussions in participants' performances. On the other hand, manners of articulation are divided into six: stop, fricative, approximant, trill, tap, and lateral (Ladefoged 2014). I will focus only on the two types involved in the experiment. Stops or plosives, like /b/, "involve closure of the articulators to obstruct the airstream. This manner of articulation can be considered in terms of nasal and oral stops" (2014). Then fricatives, like /v/, are produced with "the close approximation of two articulators, so that the airstream is partially obstructed, and a turbulent airflow is produced" (2014). Having located consonants in a specific manner will allow us to discover if their different production differs the perception of them.

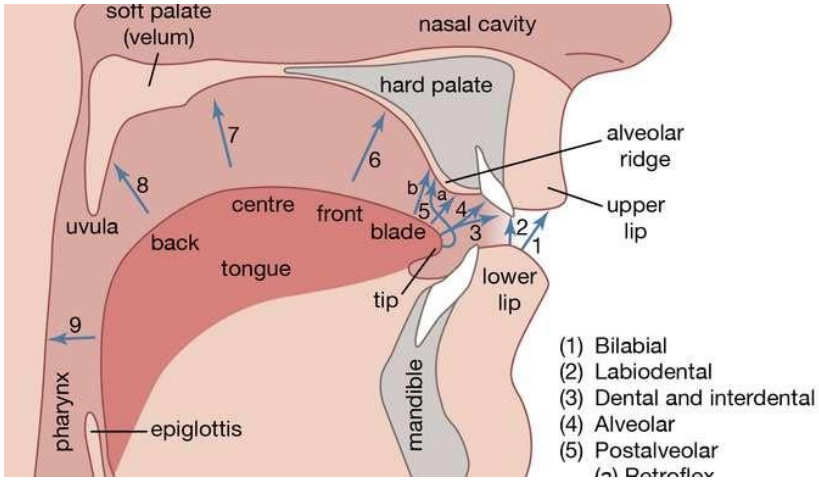

Fig.1. Elements of the vocal tract adapted from Encyclopaedia Britannica

1.4 Recent literature on the impact of masks

Although the insertion of masks is quite a new issue, some scholars have started studying the impact that masks have in communication. For instance, Louise Martin focuses on the difficulties that people with hearing loss suffer as a consequence of communicating with masks (2020,28). She lists all the aspects that are affected. Firstly, the reduction of "sound levels at high frequencies" and "amplification at low frequencies" provoking an unclear effect on understanding (2020, 28). Further on, she puts emphasis on the fact that the environments with noise, together with the need of wearing masks would affect people with hearing loss because

they would suffer most in such circumstances (2020,28). Moreover, one of the drawbacks that she mentions is directly connected with the purpose of this experiment and it is related with the loss of visual-speech cues. Auditory and visual cues allow speech perception to be productive, and people with hearing impairment depend more on the visual information, especially when they are dealing with noisy surroundings. These visual-speech cues are blocked by medical or cloth masks. Although the impact of this obstruction in perception differs taking into account the degree of hearing-loss, a negative impact is what characterises the outcome. The ability of speechreading becomes impossible for people who highly rely on it. Nevertheless, there is still a lot of research in need concerning the variability of hearing impairments that exists (2020, 28).

On another note, there are also some investigations that have been made in relation to the usage of the different types of masks, and how their effects may vary depending on the material they are constituted with. In order to help those with hearing impairments, transparent masks were created. They may be seen as beneficial because they allow listeners to see the mouth movements. Atcherson together with other researchers, have analysed the effects of transparent facial masks, given the fact that "the major problem is that typical masks (cloth or medical) present an obvious visual barrier to those who depend on nonverbal communication cues on the face (e.g., mouth, lips, teeth, tongue, and cheeks)" (Atcherson et al 2020, 21). They conducted a study comparing the Sound Pressure Level (SPL) between surgical masks and transparent masks and they found out that "although transparent masks were shown to reduce SPL and conceivably further degrade speech more than their non-transparent counterparts, they play an important role in preserving nonverbal communication cues on the face" because they "help maintain access to the mouth, which can help to aid some listeners with lipreading and other nonverbal cues, such as emotion" (Atcherson et al 2020, 25). Given this information, it is evident that masks influence communication, and everything that has been studied up to this day points towards an impossibility to find the solution to have an equality between "mask" and "no-mask" communication.

1.5 The present study

Since the pandemic situation is relatively new from a global perspective, the presence of masks has been scarcely studied in relation to the act of communication. Masks might have an effect on the listeners and their perceptual abilities, moving the act of speaking towards a more listening-related activity rather than visual-auditory speech perception. The aim of the project

is to investigate the relationship between auditory perception skills and facial gestures or mouth movements when it comes to the English perception of speech sounds. An experiment with participants from secondary school levels has been carried out to analyse the influence of masks in the understanding of a set of minimal pairs words containing the combination of the voiced labio-dental fricative sound /v/ and the voiced bilabial plosive /b/.

The experiment will try to demonstrate that the presence of masks in class affects the students' understanding of English due to the limitation of not seeing facial gestures when listening to words. Having offered an account of previous literature dealing with the study of speech perception, the development of the auditory skills together with the influence of facial gestures and how perception has been related to both auditory and visual articulators, the structure of the paper is as follows. Firstly, I will dive into the method used to carry out the experiment including participants, stimuli materials and analysis procedures. To continue, I will report an analysis of the general results containing the score of each individual participant to illustrate the possible differences that might exist between the two conditions ("mask" and "no-mask"). In the same section I will fully report an item analysis to depict the dissimilarities between the set of minimal pairs. Finally, a discussion about the implications of the experiment will be developed in which limitations and possible pedagogical applications to this problem will be offered. Apart from this, a small remark on the differences depending on the participants' language background will be carried out.

2 Method

2.1 Participants

The coverage of this experiment involves the participation of 10 different people. They were asked information regarding age and language background. They were secondary school students belonging to the two first years (1st and 2nd of ESO). Their ages range between 12 and 14 years old. According to the Common European Framework of Reference for Languages (CEFR), their level of English coincides between the A2-B1 level. As far as the mother tongue is concerned, they speak mostly Spanish (4 participants) or Catalan (4 participants) at home with the exception of one participant that has Czech as their first language and another participant which is a Catalan-Spanish bilingual. Nine of them attend English lessons outside school and none of the participants has lived abroad in an English-speaking context. All the personal data from the participants was collected in an initial online questionnaire through Google Forms and the answers are collected in a table that I created using an Excel file (Table 1).

PERSONAL QUESTIONNAIRE					
PARTICIPANTS	AGE	SCHOOL YEAR	LANGUAGE(S)	ENG. OUTSIDE	ABROAD
PARTICIPANT 1	13	2º ESO	SPANISH	NO	NO
PARTICIPANT 2	13	2ºESO	SPANISH,CATALAN	YES	NO
PARTICIPANT 3	12	1ºESO	CATALAN	YES	NO
PARTICIPANT 4	12	1ºESO	SPANISH	YES	NO
PARTICIPANT 5	12	1º ESO	SPANISH, CZECH	YES	NO
PARTICIPANT 6	12	1ºESO	CATALAN	YES	NO
PARTICIPANT 7	12	1ºESO	SPANISH	YES	NO
PARTICIPANT 8	12	1º ESO	SPANISH	YES	NO
PARTICIPANT 9	14	2ºESO	CATALAN	YES	NO
PARTICIPANT 10	12	1ºESO	CATALAN	YES	NO

Table 1. Participants' information

2.2 Stimulus Material

Due to the sanitary situation and the restrictions regarding the pandemic, the experiment was designed to be carried out in a distance environment. A minimal pair task was the method used to develop the project and to test the hypothesis. A video recording was sent to the participants together with two different Google Forms documents containing all the instructions. Technology and the use of the ICTs were essential to develop the experiment since any of the procedures was done in a face-to-face environment.

The first procedure that I followed was to send a short questionnaire that gathered basic personal information about the participants' language background as I have explained in the "*2. 1 Participants*" section. Such information was collected to see if it had a possible effect in the analysis of the results. After the 5 questions, participants were provided with the video recording containing the acoustic materials and a Google Forms answer sheet. I recorded myself in order to provide the acoustic materials for the activity. The purpose of the activity for the participants was to see the video and at the same time guess which of the minimal pair items was pronounced. For this reason, the participants were instructed to have two electronic devices to be able to reproduce the video and at the same time answer the questionnaire.

The activity itself was divided into two sections that belonged to the two different conditions that the paper covers: "mask" and "no-mask". The video had a duration of approximately a minute and words were uttered with a time-distance of two-three seconds to make sure participants were able to read and select the answers. In total, there were 10 different minimal pairs (Table 2), which were used for both sections. So, participants heard the ten first in a specific condition and the following ten in the other condition. It was decided that the participants were allowed to play the video once so that no practice effect nor repetition would have an effect on the results. In the same line, the target words were presented in a counter-balanced order. As a consequence of this, the materials were 10 different video-recordings and

10 different answer sheets, one for each participant. In Figure 3, I show an example of the test that was sent, in this case to the first participant.

1. BERRY/VERY
2. BOAT/VOTE
3. BET/VET
4. BAN/VAN
5. BEST/VEST
6. REBEL/REVEL
7. BOWELS/VOWELS
8. BOW/VOW
9. BOLT/VOLT
10. SERB/SERVE

Table 2. Minimal pairs

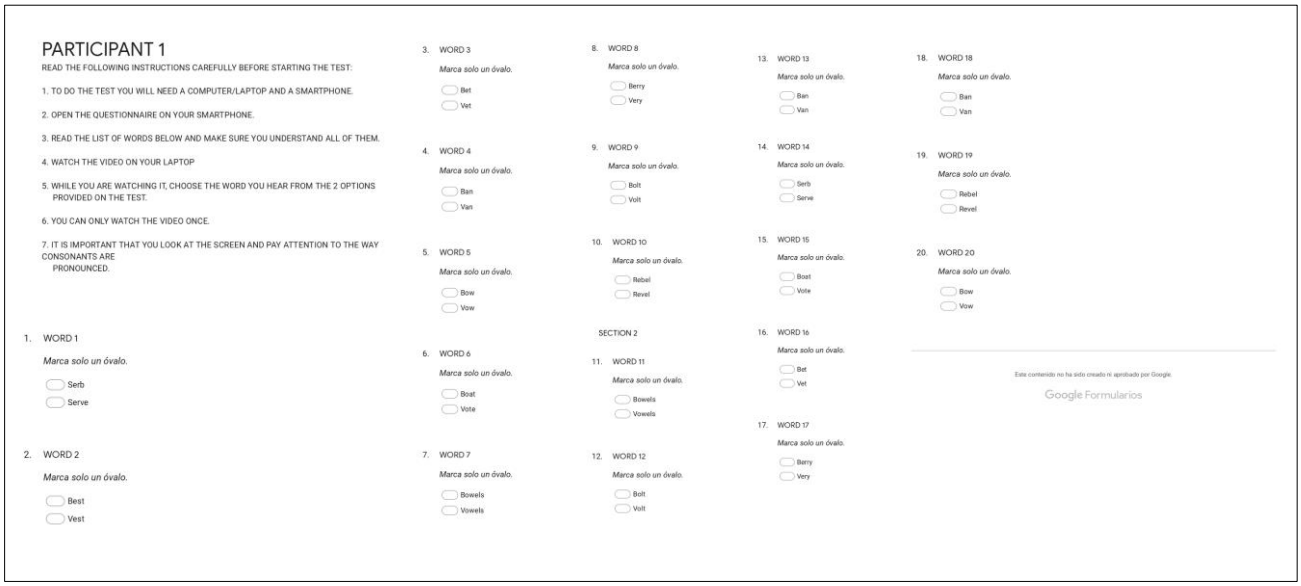

Fig. 2. Participant answer sheet

2.3 Analysis procedure

Once the answers were completed, they were extracted from the Google Forms and an Excel file was created to collect the results and to facilitate the comparisons and analysis of the experiment. I went through the answers of each participant in order to see if they recognized the correct consonants in each minimal pair, which I will refer to as hits in this paper. To continue, I gathered the hits in a table and organised the answers on the basis of "mask" and "no-mask" perceptive assimilation variables (Table 3). Therefore, the results that will be

detailed in the following section only account for the number of hits that each participant makes. The data was transported into a table using IBM SPSS in order to obtain the necessary statistical values. This program automatically extracts the mean, the standard deviation and the t-test. With this information, IBM SPSS creates graphics and tables to display the results (Table 4). The information displayed on the table was translated into English and saved as an image. Later on, I focused on the stimuli material. Another table gathered the number of hits belonging to each item of minimal pairs and a similar procedure was followed in order to illustrate the differences between participants' performance (Table 5). From both sets of data (participants' differences and minimal pairs' differences), I extracted the difference in the number of hits from the mask condition and the no-mask condition, as seen in the tables previously mentioned. Finally, to offer a more visual approach of the results, I also created histograms using Google Sheets for the number of hits that are presented on the results' section.

PARTICIPANTS	NO MASK	MASK	DIFFERENCE
PARTICIPANT 1	9	7	2
PARTICIPANT 2	10	5	5
PARTICIPANT 3	6	6	0
PARTICIPANT 4	4	8	-4
PARTICIPANT 5	4	4	0
PARTICIPANT 6	4	3	1
PARTICIPANT 7	5	6	-1
PARTICIPANT 8	9	5	4
PARTICIPANT 9	9	8	1
PARTICIPANT 10	8	3	5

Table 3. Participant's hits

Paired Sample T-test

		Paired differences					t	degrees of freedom	Sig. (bilateral)
		Mean	Standard Deviation	Average Error Deviation	95% confidence interval of the difference Inferior	Superior			
Par 1	No_Mask - Mask	1,30000	2,83039	,89505	-,72474	3,32474	1,452	9	,180

Table 4. Paired-samples t-test examining the effect of mask condition

MINIMAL PAIRS	NO MASK	MASK	DIFFERENCE
BOWELS/VOWELS	5	6	-1
BOW/VOW	6	2	4
BOLT/VOLT	6	5	1
BERRY/VERY	6	7	-1
BAN/VAN	7	2	5
BEST/VEST	7	6	1
SERB/SERVE	7	3	4
BOAT/VOTE	8	8	0
BET/VET	8	4	4
REBEL/REVEL	8	5	3

Table 5. N*umber of hits out of 10 obtained by the participan*ts in the "no-mask" and "mask"

3 Results

3.1 Overall results

The research question addressed in this study concerned whether the presence of a mask in a section of the task influenced the perception of the target words. As it was explained before, in the activity participants were exposed to language in a controlled way. A selected set of minimal pairs were reproduced, and they had to choose the one they heard. Taking a general perspective of the results, it was found that the common behaviour was that participants got a lower score in the mask condition. According to the statistical values, the mean for the ten participants under the mask condition was that of 5'5 hits while without the mask environments the mean was slightly higher, 6'80. Fig. 4 and fig. 5 show graphically the score of each participant in each condition.

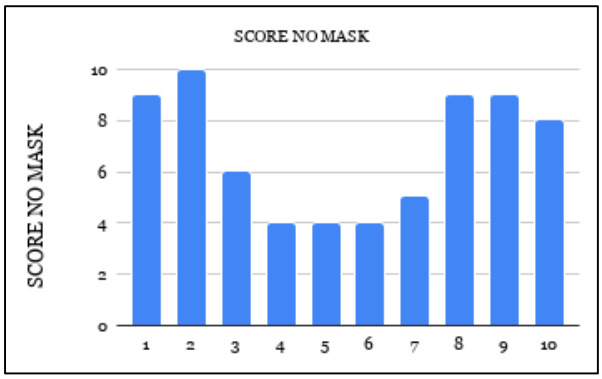

Fig. 3. Mean score no-mask

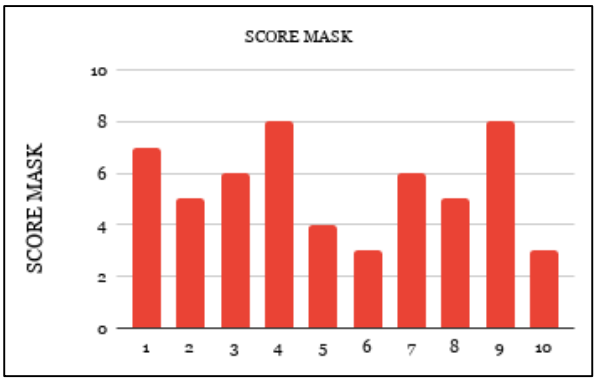

Fig. 4. Mean score with mask

Analysing the results across participants, a high degree of between-subject variability was observed. In the "no mask" situation, results vary from the highest result of 10 hits, only reached by one participant, to the lowest score of 4 hits by 3 participants. Values in the middle range correspond to 5 hits (1 participant), 6 hits (1 participant), 8 hits (1 participant) and 9 hits (3 participants). The mask condition is also unstable, the highest score belongs to two participants that made 8 hits, and the lowest was 3 hits by two participants. The other scores were 4 hits (1 participant), 5 hits (2 participants), 6 hits (2 participants) and 7 hits (1 participant). All this data reflects that there is not a constant performance in any of the conditions and the number of scores varies significantly.

Another instance of dissimilarities is within-subject variability. Generally, the majority of responses show a better score when participants listened to the words pronounced without mask, but, when it comes to the mask condition perception, some participants who had a high score in the previous section, did also receive a high score in this condition. Nevertheless, this cannot be considered the regular pattern, because there were also participants whose difference in performance is greater. The biggest difference between the number of hits were 5 words belonging to two participants, whose results were 10-5 and 8-3. The smallest difference belongs to two participants that obtained the same result in the two variables, 4-4 and 6-6.

From the participants' performance, results can be summarised as the following: six participants confirm that their perception was better in a no mask environment, two participants remain stable in the two conditions and another two participants obtained more hits in the mask condition than in the no-mask condition.

3.2 Item analysis

The specific material selected for the experiment was used in a randomized manner in order not to infer participants' responses. All the answers' sheets were presented differently, they

contained the same minimal pairs, but the sequence of pronunciation was different. Consequently, the number of times that the consonants /b/ or /v/ were pronounced underwent a process of randomization. This explains the difficulty of analysing possible reasons of variability in the responses. Nevertheless, after observing the results, some remarks can be made. The set of the ten minimal pairs was analysed in terms of the times they were understood as intended in order to determine if there were coincidences. Seven minimal pairs confirmed the hypothesis of the present study, that is, they are properly recognized in a no-mask environment. These pairs are Bet-Vet, which in the no-mask condition had a total score of 8 hits contrary to the 4 hits in the mask condition setting, Ban-Van (7 hits in "no-mask" vs 2 hits in "mask"), Best-Vest (7 vs 6), Rebel-Revel (8 vs 5), Bow-Vow (6 vs 2), Bolt-Volt (6 vs 5), and Serb-Serve (7 vs 3). On the contrary, there were two sets of minimal pairs (Berry-Very and Bowels-Vowels) that were slightly better discriminated in the mask section of the activity, with 7 and 6 hits respectively. Both exceeding in one hit the no-mask situation. There was also another pair (Boat-Vote) that was equally recognized in both environments with a total of 8 hits. Results are portrayed in figure 6.

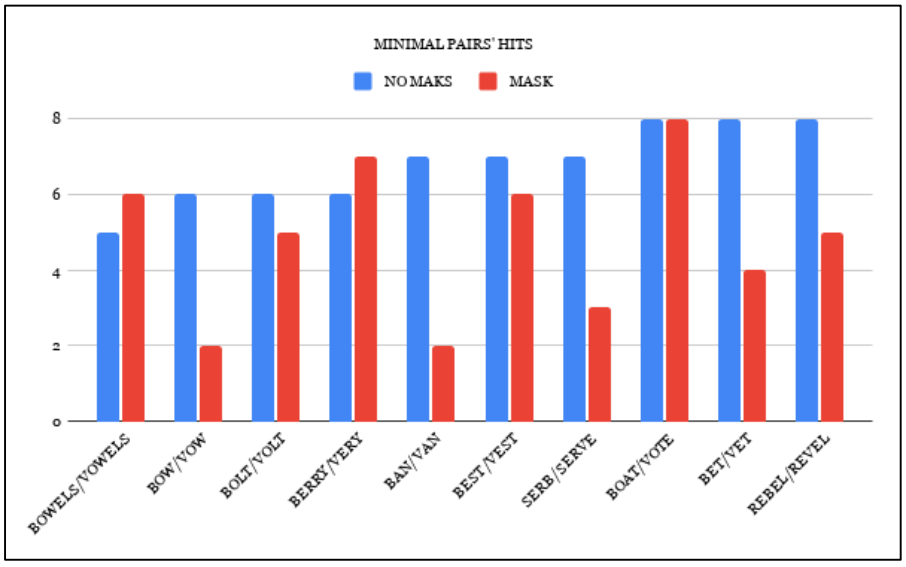

Fig. 5. Minimal pairs' hits

4 Discussion

4.1 Limitations

The development of this experiment has enabled an early investigation into the question whether the presence of masks have actually had an impact on the way students perceive

English language consonants. Nevertheless, the results should not be taken as determinant. According to the statistical value as shown previously in table 4, the paired samples t-test obtained by comparing the two conditions ("mask" and "no-mask") cannot reject the null hypothesis, this means, the hypothesis that shows no significant difference between specified populations. The reasons that possibly explain this might be related to the limitations that the experiment suffered in relation to the sample, the type of task, and the environment in which the experiment was executed. 10 participants are probably not sufficient to draw significant conclusions, and a greater sample would help to obtain more variety in the results and, therefore, greater possibilities to reject the null hypothesis. Apart from that, it is important to remark that the task was conducted in a distance environment, outside the classroom and with digital material. It is still to be seen how results would change in a face-to-face format, with participants inside the same classroom and the minimal pairs pronounced in a face-to-face setting.

4.2. Differences depending on the mother tongue

The personal information about the participants serve also as a factor that needs to be taken into account in order to relate the type of task and the analysis of the results. The focus of the activity was on the consonants /b/ and /v/, which in English show differences in terms of manner of articulation and the position in which lips and teeth are placed. As Erdener explains, "Cross-language studies of auditory–visual speech perception suggest that there appear to be interlanguage differences in the amount of visual speech use" (2013, 122). In his study, he concludes that when participants "were presented with non-native speech stimuli, the visual speech influence they used increased significantly". (Erdener 2013, 122). However, the present study shows surprising results. In the case of Spanish, the language that dominates among participants, the distinctions between the consonants /b/ and /v/ are not so significant and when it comes to pronunciation, no distinction is made. On the contrary, in the Catalan language, the same consonant characteristics as in English exist. Interestingly, results reflect that the group of Catalan-dominants, who are more used to hearing the difference in these consonants, show a worse score when they hear the minimal pairs in the mask condition, except for one participant whose performance was the same. In the case of the Spanish-dominants, who do not have this distinction in their L1, show a striking result since some participants had a greater number of hits when they heard the words with me having the mask on. Having analysed this situation, it could be said that maybe, in this specific case, the mother tongue does not have an effect on the acoustic perception of the participants.

4.3 Implications from a teaching perspective

Lastly, from the educational point of view, this situation has questioned the students' needs in an English as a Foreign Language class that can also be extended to any other subject. As mentioned before, any of the different types of masks have their drawbacks as well as benefits when it comes to achieving proper communication. When transparent masks appeared to be the solution to the facial barrier problem, they resulted to be worse in the sound pressure aspect. So, up to this moment, a perfect solution to the problem of speech perception seems to be unattainable. However, this project has made clear that specific attention to the students' necessities should be implemented. It is highly important to attempt to provide a comfortable atmosphere to everyone present in class. Following this line, Manoel Nobrega together with other scholars have elaborated a list of guidance to teachers. They offer advice that teachers can follow during the Covid-19 pandemic in reference to their teaching attitudes. Among them we find the need for slow and articulate speaking, supporting activities with visual contents, avoiding environmental noise, using a microphone, asking students for repetition of the instructions previously provided, not to speak while the students are walking and not making eye contact (Nobrega et al 2020, 2). Were those instructions followed, the education community would be able to enjoy a more pleasant environment even in the sometimes frustrating situation that the Covid-19 crisis is.

5 Conclusion

From a general perspective, though, moving statistical values aside and having a view strictly to the number of hits that the specific participants of this experiment did, it seems that the question that motivated this research is confirmed and the general behaviour is that of a better score in a no-mask environment. Results declare the existence of certain difficulties for some of the participants to discriminate the right sound when they are not able to see the articulatory gestures. The presence of masks influences to a certain degree the participants' performance. A relevant aspect about this study is that I have added data about a field of study that is new and that has been scarcely studied. Although, as I have said before, the results of this specific study are not conclusive, they have aided to portray the issues that the insertion of masks has had on the field of speech perception. Moreover, they may serve as a starting point for further research. It is true that some researchers have started to show interest in the field, massive investigation should be carried out taking into account that this problem will apparently be present for a long period of time.

Works Cited

Atcherson, Samuel R., Evan T. Finley, B. Renee McDowell, and Celeste Watson. 2020. “More Speech Degradations and Considerations in the Search for Transparent Face Coverings During the COVID-19 Pandemic.” *Audiology Today* 32, no. 6 (Nov/Dec): 20-7.

Coelho, Janaina dos Santos, Renata Christina Vieira, and Esther Mandelbaum Gonçalves Bianchini. 2019. “Interference of dentofacial deformities in the acoustic characteristics of speech sounds.” *Revista CEFAC* 21, no. 4 (September): 1-10. https://doi.org/10.1590/1982-0216/201921419118

Diehl, Randy L., Andrew J. Lotto, and Lori L. Holt. 2004. “Speech Perception.” *Annual Review of Psychology* 55 (February): 149-179. https://doi.org/10.1146/annurev.psych.55.090902.142028

Erdener, Doğu. 2012. “Basic to applied research: the benefits of audio-visual speech perception research in teaching foreign languages.” *The Language Learning Journal* 44, no. 1: 1-9. https://doi.org/10.1080/09571736.2012.724080

Erdener, Doğu, and Denis Burnham. 2013. “The relationship between auditory–visual speech perception and language-specific speech perception at the onset of reading instruction in English-speaking children.” *Journal of Experimental Child Psychology* 116: 120-138. http://dx.doi.org/10.1016/j.jecp.2013.03.003

Ladefoged, P. N. 2014. "Phonetics." Encyclopaedia Britannica, accessed May 14, 2021, https://www.britannica.com/science/phonetics.

Martin, Louise. 2020. “Impact of face masks and face shields on communication.” *Hearing Review* 27(10):28-29. https://www.hearingreview.com/hearing-products/accessories/infection-control/nal-update-impact-of-face-masks-and-face-shields-on-communication

McDonald, John. 2018. “Hearing Lips and Seeing Voices: The Origins and Development of the ‘McGurk Effect’ and Reflections on Audio–Visual Speech Perception Over the Last 40 Years.” *Multisensory Research* 31 (January): 7–18. Doi:10.1163/22134808-00002548.

Nobrega, Manoel, Renata Opice, Mariana Machado Lauletta, and Christiane Ayello Nobrega. 2020. “How face masks can affect school performance.” *International Journal of Pediatric Otorhinolaryngology* 138 (2020): 1-2. Doi: 10.1016/j.ijporl.2020.110328.