Speech Pauses among Bilingual Students of the English Language and Literature

Tomas, Laura

UNIVERSITY OF ZAGREB

FACULTY OF HUMANITIES AND SOCIAL SCIENCES

DEPARTMENT OF ENGLISH

ACADEMIC YEAR 2023/2024

Laura Tomas

Speech Pauses among Bilingual Students of the English Language and Literature

Master's thesis

Thesis advisor: prof. Irena Zovko Dinković, PhD

Zagreb, 2024

Abstract

Hesitation phenomena are an integral part of spontaneous speech, which is said to be comprised of the alternation of fluent and hesitant phases. With that in mind, the aim of this master's thesis is to provide a contrastive analysis of speech pauses in English and Croatian among bilingual students of the English language and literature. The theoretical background describes the stages of spontaneous speech, offers the classification of hesitation phenomena, proposes the possible functions of speech pauses and illustrates the types of bilingualism, which are important to differentiate for the purposes of the study. The second part of the thesis presents the study of speech pauses in English and Croatian among ten students of the graduate programs of the English language and literature at the Faculty of Humanities and Social Sciences in Zagreb. Given that the participants were native speakers of Croatian and were highly proficient in English, it was hypothesized that the number and the location of speech pauses in these two languages would be approximately the same. The analysis of speech samples in English and Croatian showed that the number of pauses was roughly the same in English and Croatian (103 and 112, respectively) and in both languages, the participants mostly paused in sentence-initial positions or before major syntactic units, such as clauses.

Keywords: spontaneous speech, speech pauses, bilingualism

Table of Contents

1. Introduction

Spontaneous speech, unlike prepared or read speech, is not entirely fluent. It is being planned and articulated in real time, thus it is disrupted by various hesitation phenomena (such as repeats, false starts, self-corrections, hesitation pauses, etc.). As a student of the English language and literature, I wanted to compare the number of speech pauses (filled and silent ones), as well as their location in English and Croatian among my fellow students of the English language and literature.

This master's thesis is divided into two parts. Firstly, I present the theoretical background and existing finding regarding the stages of spontaneous speech, the classification of hesitation phenomena, and the proposed functions of hesitation pauses in spontaneous speech. Considering that bilingualism is an important aspect of this research, I also briefly illustrate various types of bilingualism.

The second part consists of the research I conducted. Given that the participants were students of the graduate programs of the English language and literature at the Faculty of Humanities and Social Sciences, I proposed two hypotheses: the number of speech pauses in English and Croatian was roughly the same and the location of these pauses in these two languages was also roughly the same. To test the hypotheses, the participants were required to answer two questions (one in English and one in Croatian) and the questions were the same for all the participants (see section 3.3.). The results are showcased for each participant, as well as across all of the participants. Lastly, I briefly discuss the limitations of the study.

2. Theoretical background

2.1. Spontaneous speech

Spontaneous speech is far from perfect. Unlike pre-planned speech, in spontaneous speech, we have to think about what we want to say, do the lexical and grammatical planning, and articulate these thoughts in real time. Therefore, it is hardly surprising that spontaneous speech is filled with many hesitations, such as silent and filled (*um* or *uh*) pauses, editing expressions (*I mean* or *you know*), repairs, self-corrections and false starts. These hesitations seem to happen due to obstacles in speech planning (Clark and Wasow 1998). Levelt (1989), according to Harley

2005), created a model of speech production according to which the process of speech production is realized in three stages: conceptualization, formulation, and articulation.

The conceptualization stage is preverbal, meaning that we conceive an intention and select the relevant information about the message that we want to produce. This stage is further divided into two sub-stages – microplanning and macroplanning. On a broader (macroplanning) level, we elaborate our communicative goals into smaller goals and retrieve the appropriate information. Following macroplanning is the microplanning stage in which we assign the correct shape to these pieces of information and decide on informational structural properties, such as topic and focus.

The second stage is formulation, in which we transform the message that we conceptualized in the previous stage into linguistic form. It consists of two major processes: lexicalization, which refers to the selection of the words that we want to assign to our message, and syntactic planning (arranging these words so that they form a sentence). At this point, phonological encoding occurs, as well, where words are being turned into sounds. It is important that sounds are produced in the correct sequence – they also specify how the muscles of the articulatory system are to be moved.

The final stage in the speech production process is the articulation stage. Chunks of speech are retrieved from the speech buffer and the phonetic plan from the formulation stage is translated into muscle movements for the articulatory organs (tongue, lips, vocal cords, etc.). If executed correctly, this model should result in smooth and fluent speech. However, more often than not, this "perfect" model is naturally hindered by numerous hesitations:

> Henderson, Goldman-Eisler, and Skarbek (1966) proposed that there were cognitive cycles in the planning of speech. In particular, phases of highly hesitant speech alternate with phases of more fluent speech. The hesitant phases also contain more filled pauses, and more false starts than the fluent phases. It is thought that most of the planning takes part in the hesitant phase, and in the fluent phase we merely say what we have just planned in the preceding hesitant phase (Harley 2005: 403).

Concluding remarks

Spontaneous speech, being produced in real time, is peppered with many hesitations. These hesitations are connected with speech planning. Levelt's (1989) model of speech production shows that the production consists of three stages. The preverbal conceptualization stage, where the speaker decides on their communicative goals and selects the appropriate information; the formulation stage, which includes lexicalization, syntactic planning and phonological encoding; and the articulation stage, in which chunks of internal speech are taken from the speech buffer and the speech is executed by the articulators. Early on in studies on speech Henderson, Goldman-Eisler, and Skarbek (1966) proposed that spontaneous speech alternates between hesitant and fluent phases, which indicates that hesitation phenomena are a normal and expected part of the speech cycle.

2.2. Classification of hesitation phenomena

Although the focus of this thesis is on speech pauses, I will briefly describe other hesitation phenomena as well. For each of these phenomena – with the exception of pauses which will be discussed in detail later in the thesis – I provide an example from the current research. According to Rose (1998), hesitation phenomena can be classified into the following categories:

   a) false starts

This hesitation phenomenon happens when, after uttering a few words, a speaker stops in mid-sentence. The first attempt at lexicalization is discarded and the speaker may either try to alter the lexicalization, as shown in example (1), or stop altogether, thus releasing the conversational turn. Maclay and Osgood (1959) have found that false starts typically occur before content words.

(1) *Uhm*, it was my first time being in Serbia, and, to be perfectly honest, *uhm*, **I wasn't**, I'm not really sure what I was expecting.

   b) repeats

A hesitation phenomenon that includes the reiteration of a lexical item in mid-sentence position is called a repeat. Rose (1998) notes that usually only one word is repeated. Clark and

Wasow (1998) also state that repeats are one of the most frequent disfluencies in spontaneous speech.

(2) *Uhm* najbolji ste prijatelji, a pitaš za neku malu uslugu ili veću uslugu **i**, **i** dalje ne znaš jel će se to dogoditi ili neće.

c) restarts

Restarts happen when a speaker utters a few words, then abruptly returns to the beginning and repeats those same words.

(3) *Uhm*, **we went to**, **we went to** a very small village that his grandparents are from.

d) self-corrections

Self-corrections are made when a speaker utters one word, which is then followed by another word that represents a retraction of the original word.

(4) *Uhm*, *uhm* but yeah, we went there in March of 2021 and we got back to Croatia in July **in**, **of** 2021.

e) lengthenings[1]

When a speaker elongates the enunciation of a word, they create a lengthening. They may seem like a good alternative to hesitation pauses when we need extra time to plan and process the information, given that they do not disrupt the flow of speech. Williams (2023) also points out that prolongations mostly occur before function words rather than content words.

(5) **I:** had flown from Zagreb to London.

f) pauses

Goldman-Eisler (1968: 13) offers a broad classification of pauses based on their location – they can either be grammatical or non-grammatical. Grammatical pauses are semantically determined and occur at grammatical junctures, i.e. they appear in the following places, which are highly predictable:

- "natural" punctuation points.

---

[1] A prolonged syllable is marked by a colon (:).

• immediately preceding a conjunction (both coordinating and subordinating).

• before relative and interrogative pronouns.

• before all adverbial clauses of time, manner and place.

• when complete parenthetical references are made.

Non-grammatical pauses typically occur in places where they are not expected, such as:

• where a pause occurs in the middle or at the end of a phrase.

• where a pause occurs between words and phrases that are repeated.

• where a pause occurs in the middle of a verbal compound.

• where the structure of the sentence was disrupted by a reconsideration or a false start.

Similarly, Rose (1998) distinguishes between three types of pauses: articulatory pauses, respiratory pauses and hesitation pauses.

I. Articulatory pauses

Articulatory pauses are involuntary pauses that occur during the production of speech sounds, specifically, as Rose (1998) notes, the stop consonants. They typically range from 50 milliseconds to 250 milliseconds. These pauses are omnipresent and are not included in the research on hesitation phenomena.

II. Respiratory pauses

Respiratory pauses are associated with respiration, occurring when a speaker pauses in order to breathe. They are typically silent, but can occasionally be followed by "an audible voiceless hissing caused by the generation of turbulent air at various points of stricture in the vocal tract" (Dalton and Hardcastle 1977, according to Rose 1998: 7). Rose (1998) also states that respiratory pauses may usually be associated with various speech pathologies, but are excluded from our research on hesitations, as well.

III. Hesitation pauses

When it comes to hesitation pauses, we are dealing with two main types: silent or unfilled pauses and filled pauses. Silent pauses are, as Harley (2005) states, moments of silence. According to Williams (2023), they are an integral part of the speech stream when they occur at clause and sentence boundaries. On the other hand, they are seen as disfluencies when interrupting a word, phrase or clause. Regarding their durations, there seems to be a discrepancy among the authors researching silent pauses. Kendall (2009) points out that, for Kowal and O'Connell (1980), the minimal duration is considered to be 270 milliseconds, while for Goldman-Eisler (1968), it ranges from 100 milliseconds to 250 milliseconds.

Filled pauses are voiced hesitations that can be vocalized in numerous ways, such as *er* and *uhm*, or they may include lexicalized fillers such as *well*, *like*, *so*, etc. They are also usually located at major boundaries, before low-frequency words and before longer clauses, thus signaling syntactic uncertainty (Williams 2022). Eklund and Shriberg (1998) found that filled pauses typically appear in sentence-initial, rather than in sentence-medial positions. Concerning their duration, Williams states that they are typically significantly longer than silent pauses.

Regarding the location of the pauses, Rochester (1973) also states that both unfilled and filled pauses are likely to precede content words rather than function words. However, Hawkins (1971) argues that filled pauses are more frequent at phrase boundaries, as opposed to within phrases. Consequently, since phrases typically start with function words, filled pauses are more likely to appear before function words than silent pauses. The opposite is true for unfilled pauses – they more likely occur before content words and within phrases. It is important to emphasize that these locations are not fixed. Maclay and Osgood (1959: 39) say that "either type can occur in any position where the other occurs and does so frequently."

2.3. Functions of speech pauses

When researching speech pauses, we have to consider the reasons why we hesitate. Rochester (1973) writes that speech pauses could reflect various cognitive processes, the emotional state of the speaker, as well as different types of social situations. He encapsulates the findings from various research, which will be shown below.

Goldman-Eisler (1961, according to Rochester 1973) found that silent pauses increased with task difficulties, whereas filled pauses remained constant. Levin et al. (1967) came to the same conclusion – the participants produced more silent pauses when they had to explain something rather than describe it. Lay and Paivio (1970, according to Rochester 1973) similarly concluded that both types of pauses increased with the level of abstractness. In another research conducted with Reynolds (1968, according to Rochester 1973), Paivio found that people paused more frequently when they had to define abstract nouns, as opposed to concrete nouns. Lastly, Siegman and Pope (1965, 1966, according to Rochester 1973) found that filled pauses increased with situational uncertainty.

A possible connection between anxiety and pauses also emerged, but we should first make a distinction between predispositional anxiety and situational anxiety. The former refers to an individual's general level of anxiety and the latter describes the stress level specific to a situation. Research conducted by Cassotta et al. (1967) showed that predispositional anxiety was negatively correlated with pauses, meaning that highly anxious participants made fewer pauses than lowly anxious participants. When it comes to situational anxiety, the findings showed that as situational anxiety increased, pausing increased as well (Krause and Pilisuk 1961, according to Rochester 1973).

Rochester also argues that the frequency and duration of pauses could be connected to social interaction. People tend to pause more when they have to interact with another person rather than speak into a microphone. In addition, pauses tend to increase in duration among people who need more approval from their listeners. Maclay and Osgood (1959) claim that filled pauses increase in dialogues because the speaker wishes to signal that they are not finished talking, keeping the conversational turn.

Concluding remarks

Hesitation phenomena can be classified into false starts, repeats, restarts, self-corrections, lengthenings and pauses (Rose 1998). Pauses are further divided into grammatical and non-grammatical pauses. Grammatical pauses occur at grammatical junctures and are not seen as disfluencies, whereas non-grammatical pauses disrupt the flow of speech because they occur in places where they are not expected, such as within phrases (Goldman-Eisler 1968). Oftentimes

pauses reflect issues with cognitive processes but they could also be an indicator of emotional state, as well as social situations.

2.4. Bilingualism

Bilingualism is a phenomenon that occurs when people of different language groups interact with one another. It is especially hard to define due to the complexity of its nature. For example, Bloomfield defines bilingualism as "the native-like control of two languages" (1933: 56, according to Mackey 1962), but this definition could be considered outdated, at least by modern standards and everything we know today about bilingualism. It would also rule out the participants involved in this research. To avoid that, we will showcase an alternative view. François Grosjean (2010:4) defines bilinguals as "those who use two or more languages (or dialects) in their everyday lives." The inclusion of dialects implies that bilingualism does not apply only to different languages. A person would also be bilingual if they use different dialects of the same language, or if they use the same language in different modalities, for example, speech and signing.

Furthermore, the second difference between these two definitions is that, in the latter, the focus is on language use rather than on fluency. Mackey (1962: 5) shares a similar sentiment when he defines bilingualism as "the alternate use of two or more languages by the same individual. A bilingual should have acquired the skills of listening, reading, writing and speaking, but the level of competence in each skill may vary."

Considering the variations of these definitions, it is evident that we are dealing with different types of bilingualism. Back in 1953, Weinreich (according to Harley 2005) established three types of bilingualism: compound bilingualism, coordinate bilingualism, and subordinate bilingualism. The first type referred to learning two languages in a single context and both of the languages were integrated in the brain (for example, a child was simultaneously acquiring both languages from infancy). Coordinate bilingualism was the opposite of the first type – a person was learning two languages in different environments and the languages were stored separately in the brain. Subordinate bilingualism was similar to the coordinate one, but the difference was

that the learning of the second language was dependent on the knowledge of the first (dominant) language.

However, Harley (2005) argues that the borders between these types are somewhat blurry. He proposes the following categories of bilingualism: simultaneous bilingualism (in which two languages are learned at the same time), sequential bilingualism (one language is learned from infancy and the other is learned after, but still early in childhood), and late bilingualism (the second language is learned from adolescence onwards). Thus, the participants in this research, having acquired Croatian in infancy and English in early childhood, are considered sequential bilinguals. In addition, it is important to emphasize that for each of these types, one language is typically dominant and this dominance can change over time.

Concluding remarks

Bilingualism is traditionally defined as having the native-like control of two languages, but this view is abandoned in favor of a new, more flexible definition, according to which bilingualism is the alternate use of two or more languages by one individual. Nowadays, we distinguish between three main types of bilingualism: simultaneous bilingualism (both languages are learned simultaneously), sequential bilingualism (the second language is learned after the first, but still in early childhood) and late bilingualism (the second language is learned in adulthood). The participants in this research belong to the category of sequential bilinguals.

3. Research

3.1. Aim

Speech pauses are a universal language phenomenon. It is implied that if the proficiency of bilingual speakers in any two languages is roughly the same, the number of pauses in the two languages should not significantly differ. The aim of this research is to investigate if there is a significant difference in the number of speech pauses in English and Croatian, as well as to analyze the location of the pauses in these two languages. For this purpose, we defined two hypotheses:

Hypothesis 1: The number of speech pauses in English and Croatian is roughly the same.

Hypothesis 2: The location of speech pauses in English and Croatian is roughly the same.

3.2. Participants

As mentioned in the previous section, it is hypothesized that there should not exist a significant difference in the number of pauses in English and Croatian in speakers who have similar levels of proficiency in both languages. In order to account for proficiency levels, ten students of the English language and literature at the Faculty of Humanities and Social Sciences were selected to participate in this study. The criteria for selection were the following:

a) all of the participants were first and second year students of the graduate programs of the English language and literature at the Faculty of Humanities and Social Sciences in Zagreb;

b) every participant was a native speaker of Croatian and there were no simultaneous Croatian-English bilinguals among the participants;

c) none of the participants had any speech impediments that could have affected the number of speech pauses.

All of the participants were female and the mean age of the participants was 25.3. All but three participants started formally learning English in the first grade of primary school. One participant began learning English in the fourth grade and two participants began learning English in kindergarten. The age when the participants first started formally learning English did not affect the results.

In order to ensure that all of the participants had roughly the same level of proficiency in English and Croatian, they were required to take the online Cambridge English Language Assessment test, precisely the General English test[2]. The scores for all of the participants ranged from 23 to 25 (the test consisted of 25 questions), which meant that they were eligible to prepare for the C2 proficiency qualifications.

Furthermore, prior to the actual interview, the participants read and signed the Informed Consent form, which said that they were given all the information about the study, that their

---

[2] The test can be accessed at: https://www.cambridgeenglish.org/test-your-english/general-english/

participation was anonymous and voluntary, and that the results would be used only for the purposes of this thesis. All of the audio samples will be deleted after finishing the thesis and the participants could withdraw from the study at any given moment.

3.3. Data collection and method of analysis

All of the participants were asked the same two questions – one in English and one in Croatian. In English, they were asked to describe, in a few sentences, their favorite trip. In Croatian, they were asked to describe what annoyed them the most about other people. The audio samples were recorded either via Zoom or by an audio recording application on the phone.Each participant spoke for approximately 1.30 minutes. The samples were then manually transcribed. Filled pauses, being easier to detect, were also manually marked, whereas silent pauses were detected using Praat[3] – a software for speech analysis.

3.4. Results

3.4.1. Results for Participant 1

In the English sample, all of the filled pauses are unlexicalized, i.e. they contain hesitation markers. They are all found at the beginning of major syntactic units (phrases, clauses and sentences) (in this case, the beginning of a subordinate clause (example (6)) or the beginning of a prepositional phrase (example (7). Here are some examples of unlexicalized filled pauses for P1.

(6) Probably the only trip I've taken **uhm** when I went to Italy to visit my then-boyfriend who was living there.

(7) He was in a very small town * **uh** on the coast.

---

[3] The software (version 6.4.19) can be downloaded from https://www.fon.hum.uva.nl/praat/. The authors are Paul Boersma and David Weenink
Phonetic Sciences, University of Amsterdam

Unfilled or silent pauses[4], on the other hand, show more variety. Some are found at the beginnings of syntactic units (beginning of an independent clause/noun phrase), while some occur within syntactic units (within a prepositional phrase):

(8) And ***** the trip to there was very long.

(9) **Uh** he was in ***** the south of Italy, very, very far away.

In the Croatian sample, P1 uses unlexicalized filled pauses as well as lexicalized filled pauses. In examples (10) and (11), the first filled pause is a hesitation marker which occurs in sentence-initial positions (beginning of a subordinate clause; main clauses are omitted). The second filled pause in example (11) appears in the medial position, between repeated words, which could indicate that P1 is trying to retrieve the following word.

(10) **Uhm** ili kad se ljudi guraju i nemaju nimalo * poštovanja za druge ljude i njihov prostor.

(11) **Uhm** ljudi koji su * nepristojni ili bezobrazni prema: * konobarima, *prodavačicama i drugim * **uhm** drugim radnicima u uslužnim djelatnostima.

An example of a lexicalized filled pause is found in example (12), where it also appears in sentence-initial position, while a silent pause appears at the beginning of the subordinate clause. Most silent pauses in Croatian in P1 appear within subordinate clauses (in example (13) the main clause is omitted). This could also reflect problems in lexical retrieval, especially because P1 is enumerating different kinds of service workers.
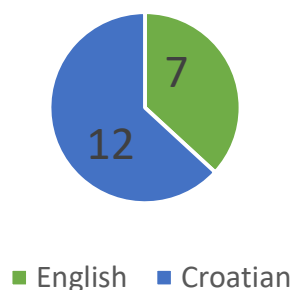
(12) **Tipa**, jedan primjer je * kad na autobusnim stanicama, na klupama ljudi sjednu na sredinu i zauzmu cijelu klupu.

(13) **Uhm** ljudi koji su ***** nepristojni ili bezobrazni prema ***** konobarima, *****prodavačicama i drugim ***** **uhm** drugim radnicima u uslužnim djelatnostima.

---

[4] Silent pauses are marked by an asterisk (*)

Chart 1. Overall number of speech pauses in English and Croatian for P1

3.4.2. Results for Participant 2

In the English sample for P2, all but two of the filled pauses are unlexicalized and each of these pauses occurs in sentence-initial position. Here are some examples of lexicalized (example (14)), as well as unlexicalized filled pauses (example (15)).

(14) **I mean**, I don't know if that counts because it was an Erasmus exchange and I was there for four months, so I don't know **if if** that's a trip or….

(15) **Uhm**, we lived there for four months, so that's a different experience than when you go only for a week.

Regarding unfilled pauses, three of the overall five silent pauses follow filled pauses and they occur in sentence-initial position, such as in example (16). In example (17), which is a compound sentence, there are three silent pauses. The first appears in sentence-initial position, between two unlexicalized filled pauses. The second silent pause appears before a second independent clause and the third pause appears before a prepositional phrase, but it notably precedes an incorrect preposition, which P2 then corrects. All of them mark the beginnings of major syntactic units.

(16) **Uhm** * well, I guess that would be my trip to Portugal.

(17) **Uhm** * **uhm** but yeah, we went there in March of 2021 * and we got back to Croatia in July * **in**, **of** 2021.

In the Croatian sample, P2 tends to use more filled pauses. Only one silent pause is found – in sentence-initial position, before an independent clause (example (18)). All of the unlexicalized

filled pauses occur before major syntactic units. In example (18), one unlexicalized filled pause appears in sentence-initial position (before an independent clause), and another unlexicalized filled pause occurs in medial position (before an independent clause, as well). In the same example, we can notice two lexicalized filled pauses, which occur within an independent clause. In example (19), another lexicalized filled pause appears within a subordinate clause.

(18) *__Uhm__, i to što su tak, __kak bih rekla__, n-nisu, __ono__, otvoreni i uvijek imaju neke pretpostavke i __uhm__ jako se striktno drže tih nekih društvenih normi.

(19) Uopće ne kuže da je to sve, __mislim__, isto ne nametnuto, ali da je to nešto što je naučeno i da se to sve može promijeniti.

This could indicate that P2 uses unlexicalized filled pauses when they are simply listing all the different things about other people that annoy them, whereas they use lexicalized filled pauses when they are trying to further elaborate on these things.
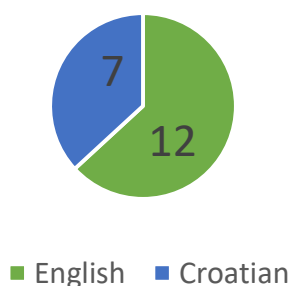
Chart 2. Overall number of speech pauses in English and Croatian for P2

3.4.3. Results for Participant 3

In the English sample, P3 prefers using unlexicalized filled pauses to lexicalized filled pauses. Only two lexicalized filled pauses can be found. The first one appears in sentence-initial position, at the beginning of a subordinate clause (example (1)), and the second one appears within a relative clause (it represents a false start), in example (2). In addition, in example (2), we can notice two unlexicalized filled pauses, which occur in medial position (within a main clause and a relative clause, respectively).

(20) __So__ since I was in Japan * at an exchange last year, I have a lot of trips that I liked there.

14

(21) I think my favorite one would be * **uhm** going to Shirakawago, which is **uhm**, it's **like**, it's protected by UNESCO.

Silent pauses occur in various positions. In example (22), a silent pause occurs in front of an independent clause. In example (23), it appears in medial position, within an adjective phrase, which could indicate the participant's effort to retrieve an unpredictable word.

(22) It has a pretty river and you can climb up and * you can see the whole: **uhm** village and those rooftops.
(23) And it's very * picturesque and scenic.

On the other hand, in the Croatian sample, P3 uses lexicalized filled pauses more often than unlexicalized filled pauses. Lexicalized filled pauses are mostly found at the beginning of independent clauses, but one also appears in sentence-final position (24) and the only unlexicalized filled pause appears within a subordinate clause, clearly indicating that P3 is trying to think of an appropriate word in Croatian (25). Examples (24) and (25) also show that in the Croatian sample for P3, silent pauses mostly appear in medial positions, within syntactic units. Only the first silent pause in example (24) occurs before a main clause, in sentence-initial position.

(24) **Mislim**, * razumijem ako je stara, starija osoba ili invalid. Ne mislim, normalno… Nego, **ono**, ak se netko * šetkara, **šta ja znam**.
(25) Ispada da ću sad imat više ovih * **uhm**, ne znam kak to reć na hrvatskom, *bigotry*, što me iritira kod ljudi.
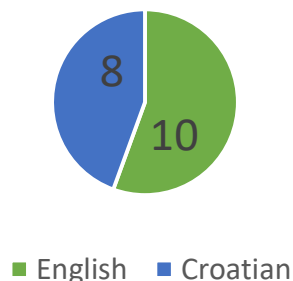
Chart 3. Overall number of speech pauses in English and Croatian for P3

3.4.4. Results for Participant 4

P4 in the English sample tends to use lexicalized filled pauses more than unlexicalized filled pauses. These lexicalized filled pauses occur mostly in front of major syntactic units, such as in sentence-initial position or before a noun phrase (26). Only one unlexicalized filled pause is found and it occurred in sentence-initial position, in front of a main clause, but it should be noted that P4 pauses after they start one sentence, but then decide to reformulate the sentence (example (27)). In examples (26) and (27), we can also notice that silent pauses appear before major syntactic units (preceding a noun phrase and a subordinate clause. One more example of a silent pause appears in sentence-initial position, as illustrated in example (28).

(26) **I mean**, the only major trip I had in my life was my senior trip in high school where we went to, * **like**, France and Spain.

(27) And then I was… **Err**, that felt really nice * because I was **like** top shit in Spain because I could actually speak Spanish and communicate with everybody and everybody else was just lost.

(28) * Favorite trip?

However, in the Croatian sample, P4 prefers using silent pauses to filled pauses. Only two filled pauses occur and both of them are lexicalized pauses that occur within subordinate clauses, as shown in examples (29) and (30). Example (29) also illustrates the use of silent pauses – they occur within subordinate clauses. To be more precise, the second and the fourth silent pause seem to indicate that the participant is still planning how to formulate the main idea, which is politics.

(29) Najviše me kod drugih ljudi živcira kad se * prave ili su uvjereni da što god oni naprave ne može imat ikakvog * političkog utjecaja * jer svi možemo utjecat na * politiku, samo se moramo, **onak**, * moramo se potrudit za to.

(30) Samo sa time da imam rozu kosu govorim da sam, **ono**, da sam drukčija od onoga što je *quote unquote* normalno.

Another example of silent pauses occurring before syntactic units (before dependent clauses) is presented in (31).

(31) Ja izađem na cestu sa rozom kosom – to je politička akcija * jer govorim * da nisam
konzerva.
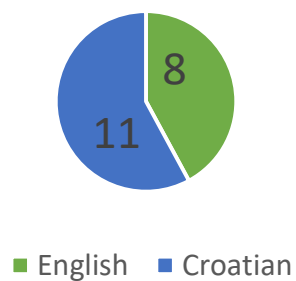
Chart 4. Overall number of speech pauses in English and Croatian for P4

3.4.5. Results for Participant 5

In the English sample for P5, filled pauses are predominantly lexicalized and only one filled pause is unlexicalized. They mostly occur in sentence-initial position (the lexicalized filled pauses in example (32) and the unlexicalized filled pause in example (33)), but the second lexicalized filled pause in example (33) occurs within a subordinate clause. These two examples demonstrate the positions of silent pauses, as well. The one in example (32) appears in medial position, but in front of a major syntactic unit (a noun phrase which follows the copula and acts as a subject complement). In example (33), however, the first three silent pauses are preceding and following filled pauses – occurring in sentence-initial position and within a subordinate clause, respectively. The fourth silent pause occurs before a main clause.

(32) **Okay**, **so** this trip was * my trip to the UK.

(33) * **Uhm**, and it was very exciting because,* **I mean** *, it's the person that you love, but also at the same time,* long-distance relationships can be a bit tricky because you never know if you will meet the love of your love or if you will get murdered.

In the Croatian sample for P5, most filled pauses are unlexicalized. Some filled pauses appear in sentence-initial positions (before a main clause), like in examples (34) and (35). In these examples, we can also notice that P5 uses silent pauses within clauses, but in example (36), two silent pauses appear in sentence-initial position, before a main clause and they are preceding and following the filled pauses, which could, in this specific example, indicate that P5 is merely emphasizing what they already said.

(34) **Uhm**, najviše me kod ljudi živcira nepoštivanje * tuđeg * mišljenja, odnosno prezentiranja.

(35) **Kao**, ti sad misliš da je toliko * tvoja misao bitnija od onoga što sugovornik govori, da jednostavno ne možeš pričekat da on završi i da nastaviš.

(36) Ili * **uhm da** *, to nepoštivanje nekako.

Some filled pauses appear within clauses (within a subordinate clause or within an independent clause), like in example (37).

(37) Specifično u razgovoru kada **uh** osoba upada * u riječ * sugovorniku ili ne da **da dovr**, da dovrši riječ, * odnosno rečenicu.
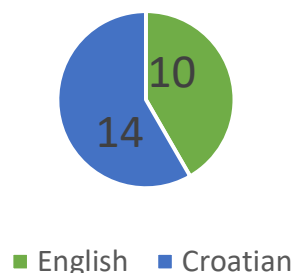
Chart 5. Overall number of speech pauses in English and Croatian for P5

3.4.6. Results for Participant 6

Most of the filled pauses in the English sample for P6 occur at major syntactic units (beginning of an independent clause or in front of a prepositional phrase), such as in example (38), but there is an example of a filled pause appearing within a subordinate clause (example (39)).

(38) **Uhm**, **so**, my favorite trip that I ever * took was * **uhm**, I think, in February last year and I went * **uh** to Dunkirk.

(39) **Uhm**, and it was * a magical thing because **uh** that was my first time to ever see the ocean.

However, the majority of all pauses in English for P6 are unfilled pauses. In example (40), some silent pauses appear in front of syntactic units (such as the first silent pause), and some appear within syntactic units, such as the second silent pause – within a prepositional phrase. The silent pauses that occur in the middle of prepositional phrases seem to indicate that P6 is trying to remember the precise details, such as the time.

(40) **So** *, I think we arrived in Dunkirk by train around * ten o'clock in the morning and then we just * walked around * the town.

In the Croatian sample, P6 uses filled pauses mostly within clauses (whether subordinate or independent), but some filled pauses appear in front of a subordinate clause, such as in example (41). Here are the examples of filled pauses for P6 in Croatian. In these examples, we can also see that for P6 in Croatian silent pauses tend to closely precede and follow filled pauses.

19

(41) Jako ne volim kada ljudi * **uhm*** imaju potrebu * **uhm*** poniziti nekoga drugoga i praviti se, **znači**, zato da bi sebe **kao** * uzvisili, da bi se osjetili inteligentnijima od druge osobe.

(42) **Uhm**, **pa sad ono**, pogotovo ako razgovaraju s nekim za koga znaju, na primjer, da ta osoba ne studira ili **ono** da nije na faksu, nego da radi neki posao i onda krene * **mislim ono kao**, već krene u razgovor s nekom pretpostavkom da je ta osoba manje * inteligentna.

Chart 6. Overall number of speech pauses in English and Croatian for P6

3.4.7. Results for Participant 7

In the English sample, filled pauses appear predominantly in sentence-initial position or in front of an independent clause, but one lexicalized filled pause appears in sentence-final position, within a subordinate clause. All but one precedes function words (mostly pronouns). Some examples of filled pauses in English for P7 can be seen below.

There are only two silent pauses in the English sample for P7 and they both appear in sentence-initial positions. One can be seen in example (44) and one appears after a false start, such as in example (45).

(43) **So** my favorite trip lasted about a month.

(44) **Uhm** * I felt… **I, I,** met… I knew only one person there, but then I met all of her friends and they all took me in just as if we had known each other for, **I don't know**, years.

(45) It was in 2019 and **uhm** I went to… * I first went to Orahovica on Ferragosto and **it's it's** a festival.

In the Croatian sample for P7, there are only two filled pauses – one lexicalized and one unlexicalized. The lexicalized filled pause appears in medial position, within a subordinate clause, and the unlexicalized filled pause appears in sentence-initial position. Here are the examples of filled pauses in Croatian for P7.

(46) Pogotovo ako ste si jako bliski, *__dakle__* neko blisko prijateljstvo, a i dalje * ako se dogovorite, vjerojatno će ili zakasniti jako puno ili uopće ne možeš računati na njih.

(47) __Uhm__, najbolji ste prijatelji, a * pitaš za neku malu uslugu ili veću uslugu __i, i__ dalje ne znaš jel će se to dogoditi ili neće.

Silent pauses in Croatian for P7 appear mostly in front of subordinate clauses.

(48) Mislim da me najviše živcira * kad je netko neodgovoran.
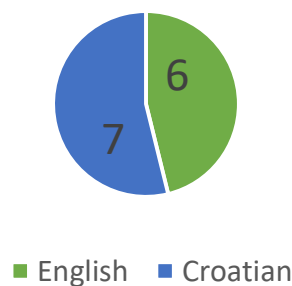
Chart 7. Overall number of speech pauses in English and Croatian for P7

3.4.8. Results for Participant 8

In English for P8, filled pauses mostly occur in sentence-initial position and/or before independent clauses, such as in example (49), but one appears within a relative clause, as in (50). In the latter example, a silent pause appears within a verb phrase, and in example (51), a silent pause is placed between two filled pauses, in front of an independent clause.

(49) __I mean__, they went above and beyond to be kind, to express their kindness.

(50) Everything that I was expecting __uhm__ kind of * fell apart because I had no idea that people in Serbia are such huge fans of Croatian people, apparently, because everyone was super, super kind.

(51) **Uhm**, it was my first time being in Serbia, and, to be perfectly honest, * **uhm**, I wasn't, I'm not really sure what I was expecting.

In the Croatian sample for P8, some filled pauses occur before independent or subordinate clauses, such as in examples (52) and (53).

(52) **Dakle** *, pokušavaju natjerati druge ljude da žive onako kako oni žele živjeti.

(53) Odnosno, **uhm** ljudi koji imaju fašistoidne neke tendencije.

Silent pauses in Croatian for P8 occur mostly within clauses.

(54) **Okej**, kod drugih ljudi me užasava kada su zadrti i kada ne razumiju da drugi ljudi * ne dijele ista uvjerenja kao oni.

(55) I ne samo da su zadrti, nego i nastoje svoj pogled na život * **uhm** proširiti na druge ljude.
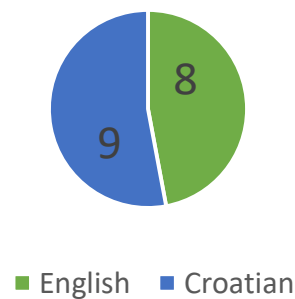
Chart 8. Overall number of speech pauses in English and Croatian for P8

3.4.9. Results for Participant 9

In English, P9 uses filled pauses mostly at the beginnings of sentences and only one filled pause is found within an independent clause. Only one filled pause is unlexicalized. Here are some examples of filled pauses for P9.

(56) **Uhm**, I really liked it because it was * very chill and I loved the landscape.

(57) I cannot remember the name of the city, but it was **like** a small city in the * center of Tuscany.

Silent pauses occur within clauses or phrases (for example, within prepositional phrases).

(58) **So** on the * first day, we just stayed in the city where the hotel was.

In the Croatian sample for P9, filled pauses appear in sentence-initial position or in front of the main clause. Only one filled pause is lexicalized.

(59) **Uhm**, **znači**, kod ostalih ljudi mi najviše smeta kada su neodgovorni i nekako nepouzdani.

(60) Na primjer, na faksu * svaki put kad je u pitanju grupni rad, **uhm** ja moram napravit najveći dio posla.

Just like the examples of silent pauses in the English sample, silent pauses in Croatian for P9 occur within clauses.

(61) **Uhm**, ispada da * ja napravim osamdeset posto posla, dok drugi naprave dvadeset posto posla, a svi na kraju dobijemo * istu ocjenu.

(62) I * to mi najviše smeta, * ta * n:eravnopravna raspodjela posla.

Chart 9. Overall number of speech pauses in English and Croatian for P9

3.4.10. Results for Participant 10

When it comes to filled pauses in the English sample for P10, we can notice some within clauses (relative clause or main clause), but there is one in front of the prepositional phrase within the main clause that we mentioned. Only one lexicalized filled pause is used and it occurs in sentence-initial position. Silent pauses also occur within these clauses. The one in example (63) appears before a prepositional phrase, and the one in example (64) appears within a prepositional phrase.

(63) And I think I **uhm** I've chosen them wisely * over the years, so every time it's great.

23

(64) But if I had to choose one, then it would be **uhm** going to Krk **uhm** in * 2022.

(65) **So**, I've been vacationing with my best friends for * over 10 years now and that year was * a bit special because finally everything * clicked into place because we were grown enough.

Most of the filled pauses in Croatian for P10 appear in sentence-initial positions, but one appears within a subordinate clause, and one appears in medial position – in front of an independent clause. In the examples below, we can notice that silent pauses are placed within clauses, indicating that the participant is still trying to formulate what they want to articulate.

(66) **Znači**, kada nisu dovoljno osjetljivi prema drugim ljudima i njihovim situacijama i kada ne stavljaju stvari u * kontekst.

(67) To jest, kada * kada im nije stalo do tuđih **uhm** osjećaja i ne razmišljaju o tome kako njihovi postupci utječu na druge ljude.

(68) Došao mi je dečko koji, za kojeg pretpostavljam da **uhm** je * negdje na na spektru **uhm** I nije znao * što je lisnato tijesto, a što je tijesto za pite.
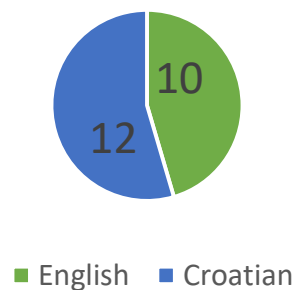
Chart 10. Overall number of speech pauses in English and Croatian for P10
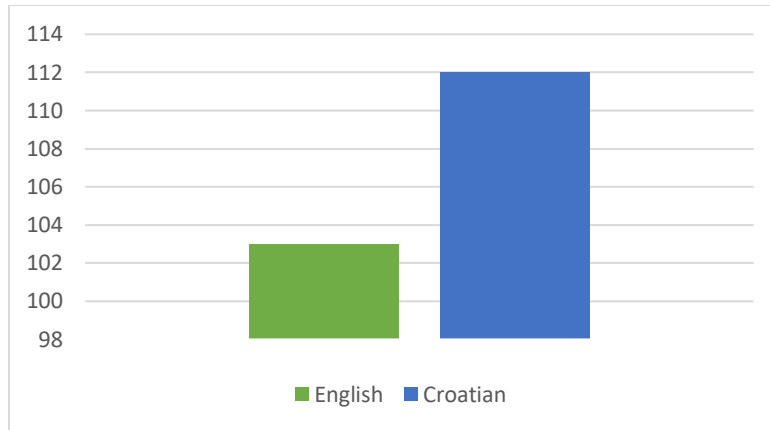
Chart 11. Overall number of speech pauses in English and Croatian across all participants.

3.5. Results across all participants in English

Chart 11 shows that there are overall 103 speech pauses in English. They mostly occur in front of major syntactic units. Specifically, filled pauses such as *so*, *well* or *okay* mostly appear in sentence-initial position, which is not unusual because this could indicate the speakers are signaling that they are about to start speaking, that they are not yet finished speaking or that they are making sure the listener is able to follow them. However, pauses within clauses could more often occur due to the speakers' ongoing (syntactic or semantic) planning or problems in lexical retrieval. Many participants showed a tendency to use silent pauses immediately before or after filled pauses. When it comes to the type of filled pauses, most of them are unlexicalized, i.e. they are hesitation markers such as *uhm* or *err*. In addition, it can be noticed that some participants had a tendency to repeat words and use self-corrections, which could also have an effect on the number of pauses.

3.6. Results across all participants in Croatian

In Croatian, the number of speech pauses is slightly higher (overall 112 recorded pauses). Just like in English, there is a tendency in many participants to start their sentences with filled pauses, thus they often occur in sentence-initial position. They can also be found in front of major syntactic units. Silent pauses are found mostly within clauses and they seem to indicate that the

25

participants are trying to remember the exact details or that they are still planning what they want to say. When appearing in clause-internal positions, filled pauses seem to indicate that the participants are trying to further elaborate on what they are saying or they are trying to reformulate their sentence.

3.7. Limitations of this research

It is important to emphasize that this study has certain limitations. The first one is the number of participants: ten participants are not a representative sample, hence the results cannot be generalized or taken as firm evidence. Secondly, in terms of syntax, English and Croatian differ significantly, thus the comparison of the location of speech pauses had to be conducted in broader terms (for example, in Croatian I only focused on the location with regard to clauses). Furthermore, the number and the location of speech pauses also may depend on the topic of the conversation. Some participants may have found one question easier to answer than the other. Individual differences among the participants should also be considered.

4. Conclusion

The aim of this study was to investigate whether there were significant differences in the number and location of speech pauses in English and Croatian among bilingual students of the English language and literature. The first hypothesis regarding the number of speech pauses in English and Croatian appears to be correct – there were 103 pauses in English and 112 pauses in Croatian, which indicates that the number of pauses was similar in both languages. Also, between individual participants, the number of pauses in each sample did not significantly differ. Most of the participants had a slightly larger number of pauses in Croatian, but I believe that this was the result of the Croatian question being more abstract – the participants had to express their opinion in Croatian, whereas in English, they merely had to retell an event that already happened. Regarding the second hypothesis, the results indicate that the location of pauses do not differ in English and Croatian – most pauses in both languages appeared in sentence-initial position or before major syntactic units. The pauses that appeared within syntactic units seemed to mostly happen when the participants were trying to remember the details or when they

encountered issues in lexical retrieval. Given the limitations of this study, it would be interesting to conduct the same research, but with a larger sample, as well as with various conversation topics.

References

Clark, Herbert H., and Thomas Wasow (1998). Repeating Words in Spontaneous Speech.*Cognitive Psychology* 37 (3): 201–42. https://doi.org/10.1006/cogp.1998.0693.

Goldman-Eisler, Frieda (1968). *Psycholinguistics: Experiments in spontaneous speech*.

Gósy, Mária (2023). Occurrences and Durations of Filled Pauses in Relation to Words and Silent Pauses in Spontaneous Speech. *Languages 8* (1): 79. https://doi.org/10.3390/languageP8010079.

Grosjean, François (2010). *Bilingual: Life and Reality*. Cambridge, Mass: Harvard University Press.

Harley, Trevor A. (2001). *The Psychology of Language: From Data to Theory*. 2nd ed. Hove: Psychology.

Hawkins, P.R. 1971. The Syntactic Location of Hesitation Pauses. *Language and Speech* 14 (3): 277–88. https://doi.org/10.1177/002383097101400308.

Kendall, Tyler S. (2009). *Speech rate, pause, and linguistic variation: An examination through the sociolinguistic archive and analysis project*. Duke University

Mackey, William F. (1962). The Description of Bilingualism. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 7 (2): 51–85. https://doi.org/10.1017/S0008413100019393.

Maclay, Howard, and Charles E. Osgood (1959). Hesitation Phenomena in Spontaneous English Speech. *WORD* 15 (1): 19–44. https://doi.org/10.1080/00437956.1959.11659682 .

Rochester, S. R. (1973). The Significance of Pauses in Spontaneous Speech. *Journal of Psycholinguistic Research* 2 (1): 51–81. https://doi.org/10.1007/BF01067111.

Rose, Ralph Leon (1998). *The communicative value of filled pauses in spontaneous speech*. MA Diss., University of Birmingham

Williams, Simon (2023). *Disfluency and Proficiency in Second Language Speech Production*. Cham: Springer International Publishing.