An attempt to identify language-universal and language-specific patterns in the use of filled pauses and prolongations

Evidence from monolingual and bilingual speakers of Russian, Hebrew, and Mandarin Chinese

Marianna Beradze,[1] Tatiana Verkhovtceva,[1] Xiaoli Sun,[2] Kristina Zaides,[3] Natalia Bogdanova-Beglarian,[4] and Natalia Meir[1]

[1] Bar-Ilan University, Israel | [2] Southeast University, China | [3] Independent Researcher | [4] S-Petersburg State University, Russia

This study examines filled pauses and prolongations in Mandarin Chinese, Russian, and Hebrew by comparing monolingual and bilingual speakers to identify both universal and language-specific disfluency patterns. Data were collected from monologues produced by monolinguals and two bilingual groups: Russian-Hebrew speakers who acquired both languages in early childhood, and Mandarin Chinese-Russian speakers who learned Russian later as a second language (L2). Analyses focused on the frequency and types of disfluencies. Monolinguals showed similar disfluency rates across languages, suggesting some universal patterns. Early bilinguals mirrored monolingual patterns in both languages, likely due to balanced early exposure. In contrast, Mandarin-Russian bilinguals exhibited higher disfluency rates in L2-Russian, likely due to increased cognitive load during speech planning. Additionally, they produced unique filled pause types not found in monolinguals, reflecting cross-linguistic transfer. These findings highlight how factors such as language proficiency, language exposure onset, and typological differences shape disfluency patterns in bilingual speech.

Keywords: hesitations, filled pauses, prolongations, monolingual, bilingual, Russian, Hebrew, Mandarin

https://doi.org/10.1075/jsls.00050.ber | Published online: 6 September 2025
Journal of Second Language Studies ISSN 2542-3835 | E-ISSN 2542-3843

1. Introduction

Disfluencies — such as silent pauses, filled pauses and prolongations[1] (sound lengthenings without tension or struggle), repetitions, and self-corrections — are universal features of spontaneous speech (Didirková, 2024; Lickley, 2015; Shriberg, 2001). However, language-specific characteristics (Betz et al., 2017; Lickley, 2015) and cross-linguistic interference (García-Amaya & Lang, 2020; Betz et al., 2021; Muhlack, 2023) can also be observed in the production of disfluencies across different languages and language pairs.

Fox Tree (2006) defined disfluencies as any phenomena causing a break in the smooth flow of spoken talk. Clark (2013, p. 1) further noted that "disfluency is any feature of an utterance that deviates from the ideal delivery of that utterance." The literature offers two complementary perspectives on disfluencies. One perspective conceptualizes them as natural byproducts of automatic "pure processes" that unfold without deliberate control. The other views disfluencies as strategic, reflecting speakers' intentional use of available options to manage communication. As Clark and Wasow (1998) argue, these views are not contradictory but rather mutually informative. Strategic uses of disfluency inherently depend on underlying processing mechanisms, which, in turn, are often recruited to support communicative intentions.

This research focuses on two specific types of disfluency — filled pauses and prolongations — referred to collectively as hesitation phenomena. Filled pauses typically involve inserting vocalized sounds like 'uh', 'um' in American English. Prolongations, on the other hand, occur when speakers stretch out or lengthen a speech sound within a word (e.g., saying 'soooo' instead of 'so'). Speakers may pause their speech and buy time by either lengthening segmental sounds or producing nonverbal filled pauses (Schettino & Eklund, 2023), both of which indicate hesitation while allowing speech to continue, rather than inserting long silence (Eklund, 2001).

Research highlights the varied functions of filled pauses in spontaneous speech. Maclay and Osgood (1959) emphasized their role in turn-taking, suggesting that speakers use them to hold the floor and indicate they have not yet finished speaking — effectively signaling: "I'm still in control — don't interrupt me!" (p. 41). Initially regarded as markers of hesitation or "points of failure," filled

1. In this study, filled pauses are defined as nonverbal vocalizations produced through vowel or nasal sounds (e.g., 'ah', 'eh', 'uh', 'em', 'um', 'en'), while prolongations are characterized by the marked lengthening of vowels within words. We exclude from our analysis verbal fillers (e.g., 'like'), discourse markers (e.g., 'oh'), word repetitions, paralinguistic elements (e.g., sighs, laughter), and other hesitation phenomena that may occur in speech.

pauses are now understood to fulfill diverse functions beyond signaling hesitation (Bogdanova-Beglarian & Baeva, 2018).

For instance, filled pauses can introduce new information (Arnold et al., 2003), mark uncertainty in question-answering contexts (Smith & Clark, 1993), and facilitate interpersonal coordination in dialogue (Bortfeld et al., 2001). While Smith and Clark (1993) highlight the role of filled pauses in signaling uncertainty, Tottie (2016) emphasizes that filled pauses often function as pragmatic markers that allow speakers to plan their utterances without necessarily indicating uncertainty or speech disruption. Pragmatic markers – defined as expressions that occur within a discourse segment but do not form part of the propositional content of the message – serve to structure discourse, manage interpersonal relations, or support communicative intentions rather than contribute to the literal meaning of the utterance (Fraser, 2006).

Kosmala and Morgenstern (2019) offer a functional classification of filled pauses, dividing them into three main categories. First is the planning function, which includes filled pauses that arise during the cognitive planning process. These may occur at the macro level (e.g., at the beginning of an utterance when organizing an entire thought) or at the micro level (e.g., mid-utterance when retrieving a specific lexical item or planning how to continue the sentence). Second is the reformulating function, in which hesitation indicates a need to revise or restructure the utterance, often involving repetition, repair, or the initiation of a new syntactic unit. The third category involves marking uncertainty – speakers use filled pauses to signal uncertainty, typically in contexts where they are unsure of the answer, and this often co-occurs with non-responses.

Although filled pauses have been extensively studied due to their high frequency in spontaneous speech and ease of detection, prolongations present challenges for detection and measurement. These difficulties often lead to their underrepresentation in annotated corpora, resulting in data sparsity (Betz et al., 2017). Like filled pauses, prolongations frequently signal hesitation; however, they tend to preserve a more fluent delivery than either filled or silent pauses (Betz & Wagner, 2016).

Importantly, prolongations allow speakers to hold the floor while continuing to speak, helping to avoid interruptions and maintain the conversational flow. As such, they function as a time-buying strategy that supports fluency and coherence in discourse (Eklund, 2001). Toward the end of a turn, prolongations may also signal the speaker's intention to retain control of the conversation and prevent others from taking the floor (Deme, 2013; Savino & Refice, 2000). Schnadt and Corley (2006) found that prolongations – the most frequent disfluency observed in their study – may indicate brief word-retrieval difficulties that are resolved during easily prolonged words like "the" or "to," helping to maintain speech flow.

Beyond this, prolongations can convey pragmatic emphasis (e.g., "Thank you soooo much"; Braver et al., 2016), and in languages like Mandarin Chinese, they may also highlight key points or express strong feedback while signaling hesitation (Lee et al., 2004). Notably, both filled pauses and prolongations share a common feature that distinguishes them from other disfluency types: they signal hesitation through vocalization and extended duration (Eklund, 2001). However, the use of these hesitation markers varies across languages. Swedish, for instance, tends to favor filled pauses over prolongations (Eklund, 2001), whereas Romance languages like Spanish and Italian more commonly use prolongations – particularly in word-final positions – with Spanish showing a stronger preference (Schettino et al., 2022; Llisterri et al., 2022). In this study, we examine how such patterns may reflect both language-universal tendencies – common across speakers regardless of language – and language-specific preferences shaped by linguistic or cultural contexts. The following subsection provides an overview of these cross-linguistic trends in the use of filled pauses and prolongations.

1.1 Filled pauses: Language-universal and language-specific aspects

The production of filled pauses exhibits both language-universal and language-specific properties. Clark and Fox Tree (2002) argued that most languages have at least one short, central-vowel filled pause, often nasalized – suggesting some phonetic commonality across languages. As an example of a language-universal property of filled pauses, Lo (2020) highlights that the duration of filled pauses tends to be consistent across languages. Specifically, filled pauses are generally longer than lexical vowels within the same language, a pattern observed in studies by Shriberg (2001) and Vasilescu and Adda-Decker (2007). Another cross-linguistic feature of filled pauses is that speakers typically produce two or more forms composed of one or two syllables, which are generally brief but may occasionally be lengthened (Clark & Fox Tree, 2002). This pattern is found across languages, suggesting a universal tendency toward short, easily produced filled pauses.

While some features of filled pauses appear universal, their form, function, and frequency vary widely by language. One of the most evident language-specific features of filled pauses is their phonetic form, which varies across languages – for example, between the central vowels used in English and the fronted vowels common in Hebrew and Russian. According to Lo (2020), many languages employ at least two types of filled pauses: UH, which consists of a vowel component only (e.g., 'uh' [ʌ] in American English, 'euh' [øː] in French), and UM, which includes a vowel followed by a nasal sound, typically the bilabial [m] (e.g., 'um' in English, 'ähm' in German).

In addition, Lo (2020) highlights cross-linguistic variation in filled pause form preferences, particularly between UH forms (vowel-only) and UM forms (vowel followed by a nasal). UH forms are preferred in Dutch (de Leeuw, 2007; Swerts, 1998), French (Torreira et al., 2010), and Nordic languages such as Norwegian, Danish, and Faroese (Wieling et al., 2016). In contrast, German speakers tend to favor UM forms (de Leeuw, 2007; Wieling et al., 2016). English exhibits both tendencies, with use influenced by social factors such as gender, socioeconomic status, and age. Furthermore, a recent study by Böttcher and Zellers (2024) found that UM form was the predominant filled pause form in American English (61%) and German (55%), while UH forms were more common among Russian speakers (69%), pointing to language-specific properties in filled pause form use. However, Muhlack et al. (2023) presented a contrasting finding for German, reporting that UH forms (e.g., [ə]) were more frequent than UM forms (e.g., [əm]). This result contradicts Böttcher and Zellers (2024) as well as earlier observations by de Leeuw (2007) and Wieling et al. (2016), who reported a typical preference for UM forms in German.

The frequency of filled pauses also appears to be language-specific, however, direct comparisons are complicated by the different units of measurement employed in the original studies. Lickley (2015) summarized findings from several studies that reported filled pause frequencies across different languages, typically measured per 100 words. Findings for English are mixed: Scottish-accented English showed a rate of 1.3 filled pauses per 100 words (Anderson et al., 1991), American English ranged from 1.6–2.2 (Shriberg, 1994) and up to 2.6 (Bortfeld et al., 2001), while Böttcher and Zellers (2024) reported a notably higher rate of 5.61 per 100 words. In Swedish, the rate varied between 3.6 and 4.4 across five corpora (Eklund, 2004). A study on Japanese reported a notably higher rate of 7.23 filled pauses per 100 words for male speakers (Maekawa, 2004). Russian showed an intermediate rate of around 2.2 filled pauses per 100 words in Podlesskaya and Kibrik (2007), but a much higher rate of 7.56 per 100 words in Böttcher and Zellers (2024).

In contrast, other studies reported frequencies per minute. For instance, de Leeuw (2007) found significantly more filled pauses per minute in Dutch speakers (10.1) compared to German speakers (6.3) and English speakers (8.0) during interviews addressing personal experiences and opinions. German speakers produced 8.67 filled pauses per minute (including clicks) and 5.7 per minute (excluding clicks) (Muhlack et al., 2023), exceeding the median of 1.73 filled pauses per minute reported by Fox Tree (2001), and Hungarian speakers produced 6.3 filled pauses per minute (Gósy, 2023).

Previous cross-linguistic comparisons of filled pauses and prolongations have been limited by methodological inconsistencies, including differences in tasks, measurement approaches, and analytical procedures. For example, studies have employed a range of tasks that vary in cognitive load, planning demands, and pragmatic features – such as map tasks (Anderson et al., 1991), conversations (Bortfeld et al., 2001), simulated phone calls (Eklund, 2004), spontaneous monologues (Maekawa, 2004; Podlesskaya & Kibrik, 2007), personal narratives (de Leeuw, 2007; Gósy, 2023), term descriptions (Muhlack et al., 2023), and video-based narratives (Böttcher & Zellers, 2024). Moreover, only a limited number of studies have examined filled pauses across multiple monolingual groups with different native languages within the same study (Böttcher & Zellers, 2024; de Leeuw, 2007; Wieling et al., 2016).

This study seeks to address a gap in the existing literature by offering a direct, within-study comparison of hesitations in monolingual speakers of Russian, Hebrew, and Mandarin Chinese, as well as in Russian-Hebrew and Mandarin Chinese-Russian bilinguals, using a consistent methodology, task, and measurement approach across all groups.

These languages exhibit substantial typological diversity: Russian, a fusional Slavic language, has rich inflectional morphology and relatively flexible word order; Hebrew, like other Semitic languages, has a rich and systematic non-concatenative morphology, in which most words are morphologically complex and derived from a tri-consonantal root morpheme and a vocalic pattern (Bick et al., 2011). Modern Mandarin Chinese morphology is characterized by a strong reliance on compounding, with minimal use of derivational and inflectional processes common in most Indo-European languages. Unlike the largely monosyllabic classical Chinese, modern Mandarin Chinese is rich in compound words that convey specific meanings (Wang et al., 2024). Phonologically and morphologically, Mandarin Chinese is a tonal language characterized by the absence of consonant clusters and a rich vowel inventory, including both monophthongs and diphthongs. In addition, each Mandarin Chinese character typically represents a single syllable (Wu & Kenstowicz, 2015).

By holding methodological variables constant, this study enables a more refined identification of language-specific patterns of hesitation. Such comparisons highlight the importance of investigating disfluencies across different languages within a unified study design to ensure greater precision and cross-linguistic comparability. In what follows, we describe the properties of filled pauses in each of the languages examined in the current study, beginning with Russian.

Filled pauses in Russian: Filled pauses are highly prevalent in Russian, second only to silent pauses (Kibrik & Podlesskaya, 2009). In Russian monologues, filled pauses occur at a rate of approximately 3.9 per 100 words, including both mixed forms (alternating sequences of vocalizations and silence) and pure forms (produced in isolation) (Kibrik & Podlesskaya, 2009). Various non-phonemic sounds are used to fill pauses in Russian, primarily [ə:] or [ə:m], [a:], [a:m], [m:] (Bogdanova-Beglarian & Baeva, 2018). Other studies have also shown that [ə:] is the most frequently produced filled pause in Russian (Verkhodanova et al., 2018).

In spontaneous spoken Hebrew, Silber-Varod (2013) reports that UH form (e.g., 'e') occurs ten times more frequently than UM form (e.g., 'em'). Since no Hebrew word is pronounced [e], listeners interpret [e] only as a filled pause (Silber-Varod et al., 2021). This contrasts with Russian, where [a] can serve as both a filled pause 'a' and the monosyllabic word for 'and' or 'but'. In their later work, Silber-Varod et al. (2021) found that during the Map Task Corpus (Azogui et al., 2015), in two conditions where the speaker was either a leader or follower, 89.5% of filled pauses were 'e' and 10.5% were 'em'. This is intriguing given that Clark and Fox Tree (2002) found in English that UM form is used before longer delays, while UH form is used before shorter ones. Since it is unlikely that Hebrew speakers only signal short delays, the distinction between 'em' and 'e' with regard to length of delays appears to be a language-specific feature.

In Mandarin Chinese, Yuan et al. (2016) primarily reported two main filled pauses, 'e' and 'en', in spoken monologues from a Mandarin proficiency test, while noting that each exhibited several phonetic variants. Based on their transcription observations, three vowel qualities ([a] as low, [e] as central, and [o] as rounded) were used as pronunciation alternatives for 'e', and two nasal consonants ([m] and [n]) for 'en'. As a result, multiple surface forms were attested, including 'em', 'en', 'e n', 'm' and 'n'. The authors noted the bilabial nasal [m] was rare, likely due to the absence of [em] as a legitimate rime (Yuan et al., 2016). In contrast, Strassel et al. (2005), analyzing Mandarin Chinese telephone speech, identified four types of filled pauses: 'em', 'en', 'eh', and 'ah'. Zhao and Jurafsky (2005), studying telephone conversations among Mandarin speakers in North America, found that 'uh' and 'mm' occur clause-initially, while the more frequent demonstratives zhege ("this") and nage ("that") – used as filled pauses – typically appear in nominal-searching contexts. Interestingly, 'en' was not mentioned in their data. Furthermore, Zhao and Jurafsky (2005) found that speakers from southern regions of China produced significantly more filled pauses, particularly 'uh', than their northern counterparts, suggesting a possible dialectal influence.

1.2 Prolongations: Language-universal and language-specific aspects

Prolongations, or lengthenings, are noticeably extended syllables that are perceived as hesitation, typically in relation to the speech rate of the surrounding segments (Betz et al., 2023). These prolonged segments in speech can be interpreted as interruptions to the usual timing of both articulation and perceptual processes, disrupting the continuity of speech (Gósy & Eklund, 2017). Prolongations exhibit both language-universal and language-specific properties. Their use is observed across languages, making them a universal feature of spontaneous speech, while the specific quality of the elongated vowel is shaped by each language's phonological system (Silber-Varod, 2013).

The frequency of prolongations also differs across languages, making it a language-specific phenomenon. For example, Betz et al. (2017) reported the following prolongation frequencies in normal speech: German (1.9%), Swedish (1.27%), Japanese (1.13%), American English (0.5%), and Mandarin Chinese (3.5%). Hungarian speakers exhibited 1.18 incidents per minute across different tasks (Gósy & Eklund, 2017), while Hebrew speakers produced 3.57 prolongations per minute during interview tasks (Silber-Varod et al., 2019).

Languages differ in the distribution of prolongations within words (initial, medial, and final segments), which can be linked to language-specific phonotactic constraints, such as permissible syllable structures (Gósy & Eklund, 2017). For instance, in Russian, initial consonant clusters are frequent and can consist of up to four consonants (Holden, 1978), whereas complex codas do occur but are less common (Kogan & Saiegh-Haddad, 2023).

Prolongations in Russian are distributed as follows: 36.2% occur in one-vowel words, 5.8% in word-initial segments, 18.8% in medial segments, and 39.1% in final segments (Bogdanova-Beglarian et al., 2023). In Hebrew nouns, most syllables (86.6%) are either CV or CVC, with triconsonantal onsets and codas appearing mainly in loanwords (Asherov & Bat-El, 2019). In the study of Silber-Varod et al. (2019), a significant majority (97.95%) of prolongations in Hebrew occurred in word-final position, with other positions being negligible. Notably, monosyllabic words such as the definite article ha- ("the") and the pronouns hu ("he") and hi ("she") were also classified as word-final prolongations.

Mandarin Chinese lacks both initial and final consonant clusters and permits only the nasals [n] and [ŋ] in word-final (coda) position, with codas often integrated with the preceding vowel (Wu & Kenstowicz, 2015). Prolongations are typically realized as the simultaneous lengthening of a vowel and its nasal coda (Lee et al., 2004). Tone may influence the occurrence of prolongations: soft tones are the easiest to prolong, followed by the first and fourth tones, while the second and third tones are less commonly prolonged (Lee et al., 2004). The distribution of prolongations in Mandarin words is approximately 4–1–95 across initial, medial, and final segments, respectively (Lee et al., 2004).

Moreover, the most frequently prolonged sound in each language appears to be language-specific. In Russian, the vowel [i] is most commonly prolonged across various speech contexts (Bogdanova-Beglarian et al., 2023; Verkhodanova et al., 2018), whereas in Hebrew, [e] was found to be the most frequently prolonged vowel during a structured interview task (Silber-Varod et al., 2019).

1.3 Hesitation phenomena in bilingual speakers: Between- and within-group comparisons

This study aims to examine filled pauses and prolongations in monolingual and bilingual speech. Specifically, we focus on two bilingual groups with contrasting language acquisition profiles. The Critical Period Hypothesis posits that early exposure to a language leads to more native-like acquisition and a critical period for language acquisition extends its effects to second language acquisition (Johnson & Newport, 1988). Under this framework, heritage language (HL) speakers are expected to resemble monolinguals in both of their languages, whereas late L2 learners are more likely to show non-native patterns of disfluency. HL speakers, exposed to both languages either simultaneously or sequentially before the onset of schooling, constitute the linguistic outcome of sustained contact between a minority HL and a dominant societal language (SL) (Montrul & Polinsky, 2021). Heritage speakers often face the unique challenge of maintaining both their dominant SL and their HL, with frequent daily opportunities for switching. These differences in language experience and use may lead to greater variability in how heritage speakers engage cognitive control strategies to manage their languages (Hernandez Santacruz et al., 2025). Nevertheless, the disfluency patterns of heritage adult speakers remain underexplored, leaving open questions about how hesitation phenomena manifest across diverse bilingual experiences.

In contrast, models of bilingual speech production and processing (e.g., Kroll & Bialystok, 2013) suggest that bilinguals, regardless of the age of onset of their languages, may diverge from monolingual patterns. Kroll and Bialystok (2013) reviewed a wide range of behavioral, imaging, and patient studies, concluding that both languages are simultaneously active to some extent when bilinguals use one of them. From a language processing perspective, evidence indicates that both language comprehension and production are influenced by the absolute and relative proficiency levels in each language. These proficiency levels are shaped by contextual factors and experience, and the resulting processing effects are bidirectional – each language influences the other (e.g., Kroll et al., 2012).

The frequency-lag hypothesis (Kroll & Gollan, 2014) suggests that bilinguals speak each language less often than monolinguals, leading to slower speech. As Fehringer and Fry (2007) note, bilinguals may show more hesitations in both languages due to the cognitive load of managing and inhibiting two linguistic systems, often resulting in slower speech in both L2 and, at times, L1.

Research investigating hesitation usage in bilingual adults has taken two main approaches: (1) comparing bilingual speakers with monolingual speakers (between-group comparisons), and (2) examining the same bilingual speakers across their two languages (within-speaker comparisons). These approaches often yield complementary insights into hesitation behavior but address different questions. Although some studies have reported on long silent pauses alongside other hesitation phenomena, the present study focuses exclusively on filled pauses and prolongations. These two types were selected because they are functionally comparable across languages, particularly in their role in managing conversational flow (e.g., floor-holding).

1.3.1 *Between-group comparisons: Bilinguals vs. monolinguals*

Research investigating hesitation usage by bilingual adults often compares the filled pauses and prolongations of bilinguals to those of monolingual speakers. It should be noted that almost all of these studies examined the L2 of bilinguals, which in some cases was the learners' second language, in others the language of highly proficient L2 speakers, and in yet others a heritage language. For instance, Smith et al. (2022) found that typically fluent Spanish-English bilingual adults – specifically HL-Spanish/L2-English speakers – produced more audibly prolonged speech sounds (prolongations) and inaudible prolongations (articulatory positions with little or no sound) in L2-English during a job interview task than previously reported among monolingual English speakers. García-Amaya and Lang (2020) examined L1-Afrikaans/L2-Spanish bilingual migrants in Patagonia, Argentina, a community dominant in their L2 and undergoing L1 attrition, and compared them with Spanish monolinguals. The study found that bilingual participants produced more silent pauses, more reformulations, and fewer filled pauses than monolinguals. Gilquin (2008) examined interviews with French learners of English and native speakers, finding that while both groups frequently used filled and silent pauses, learners overused them. In contrast, studies that found no differences in filled pause rates typically involved bilingual children – such as Spanish-English bilinguals compared to functionally monolingual peers with exposure to both languages (Bedore et al., 2006; Fiestas et al., 2005) – or small adult samples, such as native French speakers and American learners of French (Kosmala & Crible, 2022). It should be noted that comparisons of both languages of bilinguals to monolingual controls are rare, with few examples in child language research (Bedore et al., 2006; Fiestas et al., 2005) and heritage speaker studies (Böttcher & Zellers, 2024).

1.3.2 *Within-speaker comparisons: Bilinguals' L1 vs. L2*

Fewer studies have investigated how bilinguals differ in hesitation usage across their two languages. Among those that compared filled pauses and prolongations between L1 and L2, some found no substantial differences in filled pause distribution. For instance, De Boer and Heeren (2020) examined a group of bilingual female students – native Dutch speakers and advanced learners of English. This homogeneous group shared similar characteristics (e.g., age, gender, education level, L1 background, L2 proficiency, and linguistic environment). The study found that participants exhibited similar rates of filled pauses in both languages when asked to speak for two minutes on an informal topic – first in L1-Dutch, then in L2-English. Similarly, Peltonen et al. (2023) found no significant difference in filled pause frequency between L1-Finnish and L2-English among advanced adult English learners. Pletikosa and Rungrojsuwan (2018) noted similar findings in L1-Thai/L2-English speakers. Cuartero et al. (2023) also found that heritage Spanish speakers enrolled in a Spanish Heritage Language program in the U.S. exhibited similar filled pause rates in both HL-Spanish and SL-English.

However, other studies revealed that bilingual adults tend to produce more disfluencies when speaking in their non-native language (e.g., Bergmann et al., 2015; Gkalitsiou & Werle, 2023; Smith et al., 2022), particularly in more demanding tasks and contexts – such as speaking in front of unfamiliar interlocutors (Aswini et al., 2022). Götz (2019) found that filled pause usage among advanced L2-English learners varied by L1 background. Speakers of Greek and Italian exhibited the highest filled pause frequency, followed by French and Polish learners. Learners with Bulgarian, Chinese, and German backgrounds showed moderate use, while Dutch, Japanese, Spanish, and Swedish learners had the lowest occurrence.

Schüppert and Heisterkamp (2021) found that highly proficient L1-Dutch/L2-English lecturers produced fewer and shorter filled pauses in L1-Dutch than in L2-English lectures. Similarly, Muhlack (2023) observed more frequent filled pauses in L2 speech than in L1 for both L1-English/L2-Spanish and L1-Spanish/Basque/L2-English speakers. While the native L1-Spanish speakers demonstrated high proficiency in L2-English (B2 CEFR) through formal testing, the native L1-English speakers' Spanish proficiency was self-reported, making their exact proficiency level less certain – though they were enrolled in second-year university-level Spanish courses, suggesting at least an upper-intermediate command. Teng and Androsova (2022) examined hesitation phenomena in the spontaneous speech of four native Russian speakers and four L1-Mandarin Chinese/L2-Russian bilinguals. Both groups showed similar ratios of silent to filled pauses in their native languages, but filled pauses doubled when the bilinguals spoke in L2-Russian. The L1-Mandarin Chinese/L2-Russian bilinguals were university-level learners of Russian, with between 5 and 8 years of study. Common filled pauses in both languages of bilinguals included [a]-like forms (most frequent), as well as less frequent [am]-like and [m]-like forms. The filled pause [ɤ], similar in quality to [a] but produced with a more constricted mouth position, and [o] (less frequent), were specific to the bilingual speakers. Additionally, vowel-nasal sequences such as [eng] ([ɤŋ]) and [en] ([ɛn]) were observed only in the speech of the L1-Mandarin Chinese/L2-Russian bilinguals, with some individual variation in usage.

Bogdanova-Beglarian et al. (2023) found that L1-Mandarin Chinese/L2-Russian bilinguals used more prolongations in L2-Russian than in L1-Mandarin. Similarly, Felker et al. (2019) reported that bilingual adults produced more prolongations in L2-English than in L1-Dutch. Guz (2015), studying second-year English as a Foreign Language (EFL) teacher-training students – who were expected to reach upper-intermediate to advanced proficiency (B2+/C1 on the Common European Framework of Reference for Languages, CEFR) – found that their L2 speech, despite advanced proficiency, was slower and featured more frequent and longer filled and silent pauses than their L1 speech.

Some studies have observed a trend toward more filled pauses in L2, although the differences were not statistically significant. Felker et al. (2019) reported this pattern in L1-English/L2-Dutch bilinguals, while Fehringer and Fry (2007) observed a trend toward increased filled pauses in L2 compared to L1 among L1-German/L2-English and L1-English/L2-German bilinguals, although this difference did not reach statistical significance ($p = .059$). Together, these findings suggest that while some bilinguals show comparable hesitation patterns across their two languages, others exhibit differences, indicating that hesitation phenomena are shaped by a complex interplay of proficiency, language background, and contextual factors.

1.3.3 *Cross-linguistic influence and contextual factors*

Hesitation patterns are also shaped by cross-linguistic transfer and situational variables.

Cross-linguistic influence in bilinguals can be explained by the L1-transfer hypothesis, which suggests that filled pause realization varies cross-linguistically and that bilinguals often transfer their filled pauses from L1 to L2 without adaptation. For example, L2-English speakers frequently import hesitations from their L1 (e.g., French, Hebrew, Turkish, and Spanish), helping native speakers recognize them as non-native (Clark & Fox Tree, 2002). Transfer can also occur from L2 to L1. Hlavac (2011), found that the frequent use of 'um' among HL-Croatian speakers in Australia evidenced the transfer of hesitation forms from dominant Australian English to HL-Croatian.

Böttcher and Zellers (2024) showed that HL-Russian speakers with different L2 combinations exhibited higher usage of vocalic-nasal forms in Russian, which were prevalent in their L2 but less frequent in monolingual Russian speakers.

In contrast, De Boer and Heeren (2020) found evidence of adaptation rather than transfer: highly proficient L1-Dutch/L2-English speakers used the filled pause 'um' more frequently in English than in Dutch, mirroring the native British and American preference for 'um' over 'uh'. This suggests an adaptation to language-specific norms of filled pause usage.

Other studies link hesitation rates to proficiency and communicative context. Zhao (2023) analyzed L2-Russian monologic narratives produced by L1-Mandarin Chinese/L2-Russian bilinguals on the topic "How do you spend your holidays?" and found that proficiency level in L2-Russian was the strongest predictor of hesitation frequency. Böttcher and Zellers (2024) observed that filled pauses occurred more frequently in formal settings, among older speakers, and in bilingual contexts. Pletikosa and Rungrojsuwan (2018) found that lower-proficiency L1-Thai/L2-English speakers used more filled pauses than higher-proficiency speakers. Additionally, higher-proficiency L2 speakers tended to produce more filled pauses before clauses, while lower-proficiency speakers produced more within clauses.

1.4 The current study: Research questions and hypotheses

This study investigates hesitation phenomena, specifically filled pauses and prolongations, across both monolingual and bilingual speakers of typologically diverse languages. It first compares hesitation patterns in monolingual speakers of Russian, Hebrew, and Mandarin Chinese. Next, it explores how hesitation is realized and used by bilingual speakers with varying language backgrounds, differing in age of onset and language dominance. For instance, Russian-Hebrew bilinguals were exposed to both languages early in childhood, with their heritage language (Russian) being less dominant than their societal language (Hebrew). In contrast, L1-Mandarin Chinese/L2-Russian bilinguals acquired Russian later in life, making it their less dominant and less proficient language. Hesitation phenomena were examined in bilingual speakers' production in both languages.

The following research questions and hypotheses were evaluated.

RQ1: Are there differences in the use of filled pauses and prolongations among monolingual speakers of Russian, Hebrew, and Mandarin Chinese? Based on prior cross-linguistic research on hesitation phenomena (e.g., Eklund, 2001; Betz et al., 2017; Llisterri et al., 2022; Schettino et al., 2022), we formulated the following hypotheses: First, we anticipate differences in overall hesitation frequency, such that the use of filled pauses and prolongations will vary across the three languages studied — Russian, Hebrew, and Mandarin Chinese — reflecting language-specific norms. Second, we expected Hebrew speakers to prefer the filled pause 'eh', consistent with previous findings (Silber-Varod et al., 2021), and anticipated that Mandarin Chinese speakers would avoid nasal-vocalic filled pauses such as 'em', due to phonotactic constraints that prohibit word-final [m] in native vocabulary (Wu & Kenstowicz, 2015). Third, we predicted a higher frequency of word-final prolongations in Hebrew and Mandarin Chinese (Lee et al., 2004; Silber-Varod et al., 2019) compared to Russian, and more prolongations in monosyllabic, one-vowel words (Bogdanova-Beglarian et al., 2023).

RQ2: Are there differences in the use of filled pauses and prolongations between bilingual and monolingual speakers in their respective languages? Drawing on mixed findings from previous research on adult bilinguals with varying levels of L2 proficiency (e.g., heritage speakers, advanced learners, and highly proficient users), and grounded in theoretical frameworks emphasizing the increased cognitive demands of managing and inhibiting competing linguistic systems (Kroll & Gollan, 2014; Fehringer & Fry, 2007), we anticipate that L1-Mandarin Chinese/L2-Russian speakers will produce more disfluencies in their weaker L2-Russian compared to monolingual Russian speakers. This expectation is supported by Teng and Androsova's (2022) findings from university-level learners of Russian with 5–8 years of study, as well as broader evidence showing that disfluency rates tend to increase in less dominant or weaker second languages, particularly when compared to native speaker baselines (e.g., Smith et al., 2022; García-Amaya & Lang, 2020).

The prediction for HL-Russian/SL-Hebrew speakers is more complex, due to the limited number of studies comparing the HL and SL of heritage speakers to monolingual baselines. Notably, Böttcher and Zellers (2024) found consistently higher hesitation rates in both the HL and the SL of different heritage speakers compared to monolingual controls, suggesting that increased disfluency may stem from the cognitive demands involved in managing and inhibiting competing linguistic systems.

On the other hand, based on the Critical Period Hypothesis (e.g., Lenneberg, 1967; Newport, 1990), which posits that early exposure to a language facilitates native-like acquisition, we would expect early bilinguals such as the HL-Russian/SL-Hebrew heritage speakers to pattern similarly to monolingual speakers in both of their languages.

RQ3: Are there differences in the use of filled pauses and prolongations between the two languages of bilinguals (L1 vs. L2; HL vs. SL)? Previous research suggests that bilinguals who acquire their second language (L2) in adulthood tend to produce more filled pauses and prolongations in their less proficient language (Götz, 2019; Muhlack, 2023; Schüppert & Heisterkamp, 2021; Teng & Androsova, 2022). Several studies have found no significant differences in filled pause rates between bilinguals' L1 and L2, particularly among advanced speakers with similar proficiency and linguistic environments (De Boer & Heeren, 2020; Peltonen et al., 2023; Pletikosa & Rungrojsuwan, 2018; Cuartero et al., 2023). Based on the general trends for more hesitations in weaker L2 we hypothesize that L1-Mandarin Chinese/L2-Russian speakers will produce more filled pauses and prolongations in their less proficient language, L2-Russian, compared to their more dominant L1-Mandarin Chinese. For HL-Russian/SL-Hebrew bilinguals, previous findings on hesitation patterns among heritage speakers across their two languages are mixed. Böttcher and Zellers (2024) reported higher filled pause rates in HL than in SL across various language pairs. In contrast, Cuartero et al. (2023) found similar filled pause frequencies in both HL-Spanish and SL-English among heritage speakers more strongly exposed to their HL through enrollment in a Spanish Heritage Language program in the U.S. Given that our heritage speakers are more comparable to those in Böttcher and Zellers' study, we expect HL-Russian/SL-Hebrew speakers to exhibit higher frequencies of filled pauses and prolongations in their HL-Russian compared to their SL-Hebrew.

With respect to the function of filled pauses, we draw on the functional classification proposed by Kosmala and Morgenstern (2019), which distinguishes between filled pauses used for macro-planning – what we refer to as *Search and Planning* (e.g., at the beginning of an utterance when organizing a thought) – and those used for micro-planning – referred to here as *Lexical Search* (e.g., mid-utterance when retrieving a word or planning how to continue a sentence). Based on this framework, we expect HL-Russian/SL-Hebrew speakers to produce more filled pauses serving a *Search and Planning* function, consistent with findings that higher-proficiency L2 speakers tend to produce filled pauses before clauses rather than within them (Pletikosa & Rungrojsuwan, 2018). In contrast, we predict that L1-Mandarin Chinese/L2-Russian speakers will produce more filled pauses associated with *Lexical Search*, reflecting patterns observed in lower-proficiency L2 speakers who tend to produce filled pauses within clauses more frequently than between them (Pletikosa & Rungrojsuwan, 2018).

2. Methodology

The data, R script, and output analysis for the current study can be retrieved from the OSF repository (see https://osf.io/afrc8/?view_only=63bc6503b0754fadbacf246150cc9630).

2.1 Participants

A total of 70 elicited monologues were recorded from five groups of speakers ($N=50$). Each group was gender-balanced (5 male, 5 female). Two groups consisted of bilingual speakers, each tested in both of their languages, and three were monolingual control groups, one for each of the study's target languages.

Monolingual Control Groups. The monolingual Russian group (MoRU, $N=10$) consisted of students at Saint Petersburg State University, aged 20–21 years ($M=20$, $SD=0.4$). The monolingual Hebrew group (MoHE, $N=10$) included lifelong residents of Israel, aged 18–30 years ($M=24$, $SD=4$), recruited through personal contacts of the research team. The monolingual Mandarin Chinese group (MoCH, $N=10$) comprised students from various universities across China (e.g., Sichuan University, Guangxi Normal University, Heilongjiang University), aged 20–26 years ($M=23$, $SD=2$). All participants identified as Mandarin monolingual speakers.

Bilingual Experimental Groups. The HL-Russian/SL-Hebrew bilingual group (biRUHE, $N=10$) included participants from Russian-speaking families residing in Israel. They were recruited via personal contacts using a snowball sampling technique and ranged in age from 18 to 27 years ($M=24$, $SD=3$). All had been exposed to spoken Russian from birth through naturalistic input. However, most were less dominant in their HL — typically acquired in early childhood but unsupported in public domains — than in the SL, which is dominant in their environment (Rothman 2009). Exposure to SL-Hebrew began before the age of 5 ($M=3$, $SD=1$). Most participants reported both HL-Russian and SL-Hebrew as their mother tongues and were second-generation immigrants fluent in both speaking and understanding HL-Russian.

The L2-Russian/L1-Mandarin Chinese bilingual group (biCHRU, $N=10$) consisted of students recruited in the Russian Federation, aged 23 to 27 years ($M=25$, $SD=1$). They had studied Russian as a second language for 4 to 7 years ($M=5.5$, $SD=1$), either in China or in Russia. The level of Russian as a foreign language proficiency (TORFL) was also taken into account, with 5 speakers at the intermediate level and 5 at the upper-intermediate level (Saint Petersburg State University, n.d.). Three participants had studied Russian exclusively in China, three in Russia, and the remainder in both countries. While the peak of Russian language study in China occurred between 1960 and 1980, interest in the language is increasing again, with over 150 universities in China now offering Russian programs (Gulimire & Titkova, 2018; Liu, 2011, as cited in Qi, 2021).

2.2 Materials and data collection

To elicit monologic speech, we used *Hair Loss Treatment*, a 13-panel wordless comic strip by Herluf Bidstrup (Bidstrup, 2008). Participants were instructed to describe the sequence of a wordless comic strip while viewing it. They reviewed the entire strip prior to the task and had it in view throughout the narrative, either physically or via Zoom. Recordings were made using a digital voice recorder or through Zoom-based sessions, and all recordings were subsequently transferred to a secure computer.

Bidstrup's comics have previously been used in studies involving neurotypical individuals (Bogdanova-Beglarian et al., 2022) as well as patients with aphasia (e.g., Akhutina, 2007). This picture-based narrative task was chosen for its balance between spontaneity and structure, sitting between free storytelling and text retelling on the monologic complexity continuum (Bogdanova-Beglarian et al., 2013).

Russian-language data from L1 and L2 speakers were drawn from the Balanced Annotated Text Library, a corpus developed at Saint Petersburg State University (Bogdanova-Beglarian et al., 2013). Hebrew and HL-Russian/SL-Hebrew bilingual data were collected between February and April 2023 via Zoom due to participants' preference for remote interaction post-COVID. Linguistics PhD students conducted the sessions.

2.3 Transcription and annotation

All audio recordings of the monologues were transcribed by trained annotators. Four trained annotators independently annotated 70 monologues, with each annotator working on monologues in a language that was either their L1 or a native-like L2. Prosodic boundaries were primarily identified based on silent pauses, which were perceptually annotated and verified using PRAAT when necessary. Monologue duration was measured from speech onset to offset using PRAAT.

Russian monologues were transcribed in Cyrillic script and annotated according to the guidelines of the Balanced Annotated Text Library. Russian examples presented in this study were transliterated following the COST 2002 standards. Hebrew narratives were transcribed using a Hebrew-adapted version of the SAMPA system (Silber-Varod, 2011), while Mandarin Chinese narratives were transcribed using simplified Pinyin without tone markers. As noted by Betz et al. (2023), the annotation of silent pauses and prolongations is generally carried out perceptually, as there is currently no standardized or fully objective method for their automatic detection.

Following the approach of Thurber and Tager-Flusberg (1993), we coded silent pauses, filled pauses, and prolongations simultaneously during monologues transcription. In some cases, the presence of short pauses – typically around or below the 250 ms threshold – was verified using PRAAT (Boersma & Weenink 2019). Many prolongations that raised doubts due to potential confusion with accentuation lengthening were verified using PRAAT, and all were found to exceed 220–230 ms in duration. To ensure coding reliability, all monologues were initially transcribed and annotated by a primary researcher. Subsequently, the transcripts underwent a systematic review in joint sessions involving two coders. In instances of disagreement, a third coder was consulted, and all discrepancies regarding the identification and placement of silent pauses, filled pauses, and prolongations were resolved through consensus among the three coders.

While prosodic boundaries can be marked by a range of acoustic features – such as final syllable lengthening, pitch reset, and changes in speech rate – silent pauses have consistently been shown to be salient and reliable indicators across languages (Hilton et al., 2019; Kentner et al., 2023). For this reason, we adopted a pause-based strategy, which ensured consistent annotation and facilitated cross-linguistic comparability, particularly among languages that differ in their use of pitch or durational cues. Strong boundaries are typically realized with a silent pause, and pauses at strong boundaries tend to be longer than those at weaker boundaries (Kentner et al., 2023). However, a recent study examining seven languages – including Mandarin Chinese, one of the languages analyzed in the present study – found no robust cross-linguistic binary distinction between main clause boundaries and other syntactic contexts, nor between clausal and non-clausal boundaries (Peck & Becker, 2024). Based on this, we analyzed filled pauses and prolongations at all prosodic boundaries without differentiating between stronger and weaker boundaries – typically associated with longer and shorter pauses – although such distinctions were noted in the transcriptions. For each instance of a filled pause, annotators marked its phonological form – whether it consisted of a vowel, consonant, or a combination of both – the specific segmental realization, and its syntactic position within the utterance (initial, medial, or final).

Prolongations were annotated based on the position of the prolonged segment within the word (initial, medial, or final), following Gósy and Eklund (2017), Lee et al. (2004), and Silber-Varod et al. (2019). Consistent with Gósy and Eklund (2017), one-vowel Russian words were analyzed separately as a vowel-like category. Similarly, Hebrew one-syllable words beginning with [h] were also classified as vowel-like, rather than as word-final prolongations (cf. Silber-Varod et al., 2019), due to the inaudibility of the initial [h]. In addition, we annotated the position of prolongations within the utterance, using criteria similar to those of Lee et al. (2004). However, in our study, the initial, medial, and final positions within the utterance were determined relative to all prosodic units, without differentiating between types of prosodic units.

2.4 Functional annotation of hesitations

Filled pauses – defined as nonverbal vocalizations or nasalizations functioning as hesitation markers – and *prolongations* – marked by an extended duration of segmental material (Vitale, et al., 2023) – were annotated manually (see Table 1). Filled pauses were annotated according to their discourse functions, which were grouped into three categories: *Search and Planning, Lexical Search*, and *Search and Repair*. In our analysis, we adopted and adapted the functional classification of filled pauses proposed by Kosmala and Morgenstern (2019), who identify three primary functions: Planning (micro- and macro-levels), Reformulating, and Marking Uncertainty. While our approach is inspired by their typology, we found it necessary to refine and reframe the categories based on the specific discourse contexts observed in our data. In particular:

1. *Search and Planning* – This category overlaps with Kosmala and Morgenstern's *macro-level* planning function but is more specifically defined in our framework as occurring at the beginning of a monologue or at prosodic boundaries, where hesitation serves to fill the time required for planning the next segment.
2. *Lexical Search* – Corresponding to their *micro-level* planning function, this category is treated as distinct in our analysis. It refers to filled pauses that occur mid-utterance, typically when the speaker searches for a precise lexical item to fit the intended meaning.
3. *Search and Repair* – This function partially overlaps with their reformulation category, but we define it more narrowly: the filled pause typically follows a word cut-off and precedes a repair, highlighting the role of hesitation in facilitating lexical or syntactic correction.

Unlike *Kosmala* and Morgenstern (2019), we did not identify marking uncertainty function as relevant in our data, as our participants did not produce question-answer sequences where this type of hesitation typically occurs. Therefore, we excluded this category from our classification.

Table 1. Examples of monologue-text annotations

	Example	Filled pause				
		Code	**Type**	**Form**	**Function**	**Position**
1	*yemu prihoditsya () **e-n** n… narezat' () **e-n*** he has to () **e-n** s… slice () **e-n** *rezat' volosy / no **n** volosy eshchyo () **n*** cut hair / but **n** hair still () **n** *rastyo… rastyot.* grow… grows "He has to () **e-n** s… slice () **e-n** cut his hair, but **n** his hair still () **n** grow… grows."	6Ch-R	V	e-n	S&R	M
2	*/ i: pok… **m-m** ukazyvaet na to chto on* / and poi… **m-m** shows.3SG on that that he *lysyi.* bald "/ and he is poi… **m-m** showing that he's bald."	MoRU	C	m-m	S&R	M
3	*/ nu i on ves' takoi radostnyi idyot **e-e*** / well and he all so happy goes.3SG **e-e** *vidimo domoi//* apparently home "/ well, and he's all happy, going, **e-e**, apparently, home."	MoRU	V	e-e	LexS	M
4	*// **e-e** ve:… ve az ha- ish () mitorer ba-* // **e-e** ve:… ve az ha- ish () mitorer ba- *boker /* boker / "// **e-e** and… and then the man () wakes up in the morning. /"	SL-HE	V	e-e	S&P	S

Note. Each example is annotated using Interlinear Morpheme-by-Morpheme Glossing (IMMG) following the Leipzig Glossing Rules. Sentences are divided into italicised segments, with each segment followed immediately by its corresponding gloss line. A complete fluent free translation is provided at the end of the example. Group refers to the participant number, followed by the speaker's group, and – where applicable for bilinguals – the language of the example (e.g., 1R = participant 1, Russian monolingual; 6Ch-R = participant 6, L1-Mandarin Chinese/L2-Russian, Russian example; 6Heb-Heb2 = participant 6, HL-Russian/SL-Hebrew, Hebrew example). Type indicates the phonological type of the filled pause: *V* for a vowel, *C* for a consonant, and *VC* for a vowel-consonant combination. Form specifies the specific realization of the filled pause (e.g., *e-e, m-m, em, en*).
Function refers to the communicative function of the filled pause: *S&P* (Search and Planning), *LexS* (Lexical Search), and *S&R* (Search and Repair). Position indicates the location of the filled pause with regard to utterance: *S* for utterance-initial, *M* for utterance-medial, and *F* for utterance-final.
A double forward slash (//) indicates the beginning of a new utterance, while a single forward slash (/) indicates a prosodic break. Parentheses (( )) indicate a brief silent pause, and ellipses (…) represent a prolonged silent pause within prosodic boundaries.

2.5 Monologue general measures

First, we present the data for the monologues in terms of their length, measured by tokens, syllables, and seconds, as well as the speech rate across the five groups (see Table 2). The results indicated significant group differences in all measures of length and speech rate, as determined by one-way ANOVAs (see Table 2). Follow-up pairwise comparisons revealed that the MoRU group produced significantly longer monologues in terms of both tokens and syllables compared to the other groups. The decreased number of tokens and syllables in the monologues of the MoCH and biCHRU groups in their L1 may be attributed to the linguistic characteristics of Mandarin Chinese, while the reduced measures in the biRUHE group's HL-Russian monologues could be explained by their limited use and exposure to their weaker language. However, the biRUHE group performed similarly to the MoHE group in Hebrew, their dominant language. Notably, the biCHRU group's L2-Russian monologues were the longest in terms of duration (seconds) among all the groups.

3. Results

3.1 Filled pause and prolongation ratios

Since significant between-group differences emerged in monologue length, we calculated filled pause and prolongation ratios as a function of the number of tokens (i.e., number of words) to control for variations in monologue length. Individual differences in the use of filled pauses and prolongations were observed across participants in all groups (see Figure 1).

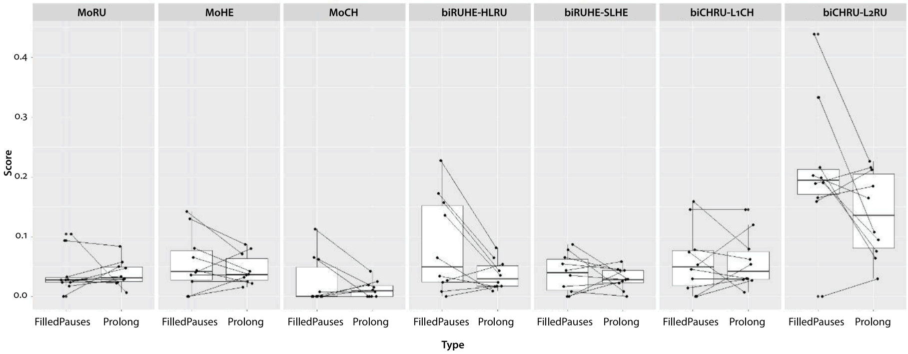

Figure 1. Filled pauses and prolongations per group (per Token)

Table 2. Length indices (M(SD) MIN-MAX) of monologues per Group

	MoRU	**MoHE**	**MoCH**	**HL–RU**	**SL–HE**	**L1–CH**	**L2–RU**	**Group differences**
Tokens (without hesitations)	210 (151)	121 (63)	66 (26)	98 (48)	123 (41)	91 (46)	141 (67)	$F(6, 63) = 3.94$, $p = 0.0021$
	42–536	62–260	40–125	37–199	46–178	48–203	79–317	
Duration in seconds (with hesitations)	100 (44)	59 (27)	57 (20)	83 (55)	55 (18)	106 (67)	178 (92)	$F(6, 63) = 6.83$, $p < .001$
	32–185	30–110	36–91	25–177	20–77	46–243	81–400	
Syllables (without hesitations)	423 (286)	199 (111)	192 (72)	182 (79)	207 (71)	269 (138)	282(153)	$F(6, 63) = 3.32$, $p = 0.00666$
	95–1034	100–455	124–351	74–345	76–327	158–613	164–697	

Note. Data are presented as mean (SD) with range in parentheses. Group differences were assessed using ANOVA. Significant results are indicated with F-values and p-values.

A linear model was fitted to predict the hesitation ratio based on the interaction between Group (with MoRU set as the reference group) and Hesitation Type (with filled pauses set as the reference type). The estimated coefficients, standard errors (SE), t-values, and p-values for the model are presented in Table 3. The linear model explained approximately 51% of the variance in filled pause and prolongation frequency, indicating that while group and type effects accounted for a substantial portion of the variance, a considerable amount of variability remains, likely reflecting speaker-specific individual differences.

Table 3. Estimates for the statistical Analyses of filled pause and prolongation ratios

Coefficients	Estimate	SE	t-value	Pr(>|t|)
(Intercept)	0.04	0.02	2.32	0.0217 *
GroupMoHE	0.01	0.02	0.28	0.78
GroupMoCH	−0.03	0.02	−1.12	0.26
GroupbiRUHE-HLRU	0.00	0.02	−0.09	0.93
GroupbiRUHE-SLHE	−0.01	0.02	−0.34	0.74
GroupbiCHRU-L1CH	0.02	0.02	0.86	0.39
GroupbiCHRU-L2RU	0.10	0.02	4.24	< 0.001 ***
HesTypeFilledP	0.00	0.02	−0.01	0.99
GroupMoHE: HesTypeFilledP	0.01	0.03	0.36	0.72
GroupMoCH: HesTypeFilledP	0.01	0.03	0.39	0.69
GroupbiRUHE-HLRU: HesTypeFilledP	0.05	0.03	1.48	0.14
GroupbiRUHE-SLHE: HesTypeFilledP	0.01	0.03	0.26	0.79
GroupbiCHRU-L1CH: HesTypeFilledP	0.00	0.03	0.05	0.96
GroupbiCHRU-L2RU: HesTypeFilledP	0.07	0.03	2.18	0.0312 *

Model fit statistics: Multiple $R^2 = 0.51$; Adjusted $R^2 = 0.46$
Note. Only significant results are presented in the table. Significance codes:
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

No significant differences were observed between the MoRU group and the other groups, nor were there differences between the two types of hesitation. However, significant interactions emerged between Group and Hesitation Type. Further pairwise comparisons using the emmeans function revealed no group differences in the ratio of filled pauses and prolongations within the monolingual groups. The biCHRU group, however, exhibited a significantly higher ratio of both filled pauses and prolongations compared to the MoRU controls (both comparisons at $p < .001$). Additionally, differences were observed between the two languages of the biCHRU group (L1-Mandarin Chinese and L2-Russian). Interestingly, no differences were found between the biRUHE group and the two control groups, nor between the two languages (HL and SL) within this bilingual group. For further details, please refer to the supplementary materials.

3.2 Filled pause realization per group

The data on different types of filled pauses per group are presented in Table 4 and Figure 2. Overall, across all monolingual and bilingual groups and languages, 'e' emerged as the most common filled pause, indicating a shared preference for this hesitation marker regardless of linguistic background or dominance.

However, Mandarin Chinese-Russian bilinguals exhibited greater variation in filled pauses when speaking in L2-Russian. While the most frequent filled pauses in this group were consistent with those observed in the monolingual Russian and Mandarin Chinese controls (i.e., 'e', 'a', 'm', 'em', 'en'), Mandarin Chinese–Russian bilinguals produced several filled pause forms not attested in the monolingual Russian or Mandarin Chinese control groups, such as 'y', 'n-n', 'an', and 'am', which appeared only rarely in the data. Notably, filled pauses such as 'en', commonly used in Mandarin Chinese, also appeared in the bilinguals' L2-Russian narratives.

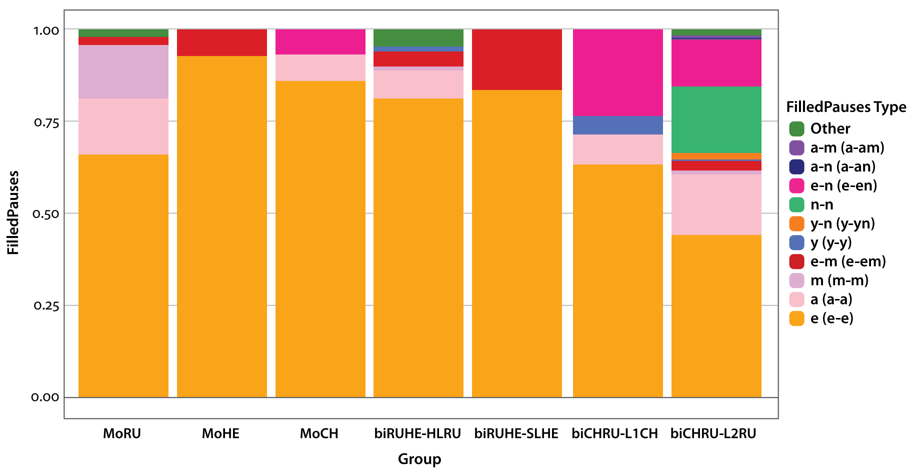

Figure 2. The visualization for the distribution of filled pause types per group

Subsequently, a linear model was fitted to predict the filled pause realization ratio based on the interaction between Group and Filled Pause. The model explained 65% of the variance in filled pause realization ratios and the overall fit was highly significant ($p < .001$). Importantly, individual differences among participants contributed substantially to the variability in filled pause use across groups and pause types. While the model accounts for a large portion of the variance through Group and Filled Pause Type effects, approximately 35% of the variance remains unexplained, reflecting considerable individual speaker variability. The estimated coefficients, standard errors (SE), t-values, and p-values for the model analyzing the distribution of Filled Pause Types are presented in Table 4.

Table 4. The output model for the analyses of filled pause type per Group

Predictor	Estimate	Std. Error	t value	p value
(Intercept)	0.47	0.04	11.13	< 0.001 ***
VocType e-n (e:n)	0.19	0.06	3.10	0.00203 **
VocType e (e:)	0.52	0.06	8.57	< 0.001 ***
GroupbiRUHE-HLRU (e:n)	−0.19	0.09	−2.19	0.02884 *
GroupbiRUHE-SLHE (e:)	−0.19	0.09	−2.19	0.02884 *
GroupMoHE (e:n)	−0.19	0.09	−2.19	0.02884 *
GroupMoRU (e:n)	−0.19	0.09	−2.19	0.02884 *
GroupbiRUHE-HLRU (e:)	0.22	0.09	2.63	0.00870 **
GroupbiRUHE-SLHE (e:)	0.25	0.09	2.94	0.00340 **
GroupMoCH (e:)	−0.20	0.09	−2.34	0.01951 *
GroupMoHE (e:)	0.21	0.09	2.41	0.01605 *

Model fit statistics: Multiple $R^2 = 0.65$; Adjusted $R^2 = 0.61$
Note. Only significant results are presented in the table. Significance codes:
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

The results of the post-hoc pairwise comparisons with the Bonferroni adjustments revealed that for most types of filled pauses (e.g., 'a-m', 'a-an', 'e-em', 'm-m', 'y-n'), no significant differences were observed between the groups. This suggests that the realization of these filled pause types did not differ between the groups. However, significant differences emerged for specific types, particularly for 'e-en' and 'e-e'. For 'e-en', statistically significant differences were found between the biCHRU-L1CH group in L2-Russian and the MoRU controls ($p = 0.042$). These findings indicate that while most hesitation phenomena were consistent across groups, certain filled pause types ('e-en' and 'e-e') exhibited group-specific variations, particularly between the bilinguals in L2-Russian and monolingual speakers. Interestingly, no differences were observed between HL-Russian speakers and monolingual Russian-speaking controls in the realization of these filled pause types.

3.3 Filled pause function per group

The data on filled pause functions are presented in Table 5 and Figure 3.

To investigate filled pause functions across the groups, we fitted a linear model that included the interaction between group and filled pause function (i.e., *Search and Planning, Search and Repair*, and *Lexical Search*). Notably, the model accounted for approximately 65% of the variance, with the remaining variation likely reflecting meaningful individual differences in filled pause use across participants. The estimated coefficients, standard errors (SE), t-values, and p-values for the model analyzing the distribution of filled pause types are presented in Table 5.

Table 5. The output model for the analyses of filled pause functions

Predictor	Estimate	SE	t-value	p-value
Intercept	0.29	0.08	3.54	< 0.001 ***
GroupbiRUHE-HLRU	0.29	0.12	2.50	0.013 *
GroupMoCH	−0.24	0.12	−2.07	0.040 *
FunctionTypeSearch&Repair	−0.29	0.12	−2.50	0.013 *
GroupbiCHRU-L2RU &Planning	−0.34	0.16	−2.06	0.041 *
GroupbiRUHE-HLRU &Planning	−0.61	0.16	−3.74	< 0.001 ***

Model fit statistics: Multiple $R^2 = 0.65$; Adjusted $R^2 = 0.61$
Note. Only significant results are presented in the table. Significance codes:
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

The regression analysis results indicate an effect of group and a significant group*function interaction. However, further post-hoc pairwise comparisons showed no significant group differences for any of the functions. Figure 3 summarizes proportional distribution of filled pause functions across speaker groups.

The most common functions of filled pauses across all groups were *Search and Planning* and *Lexical Search*, while *Search and Repair* was less frequent. Filled pauses related to the *Search and Planning* function were typically observed at the onset of a monologue or at transitional points between utterances. *Lexical Search* pauses often occurred mid-utterance, when the speaker was likely attempting to retrieve a specific word or expression. Filled pauses associated with *Search and Repair* appeared when speakers hesitated due to uncertainty about the phonetic or grammatical form of a word. This type of filled pause function was particularly common among bilingual and heritage speakers when speaking their L2 or heritage language. Since utterance position served as a foundational cue for

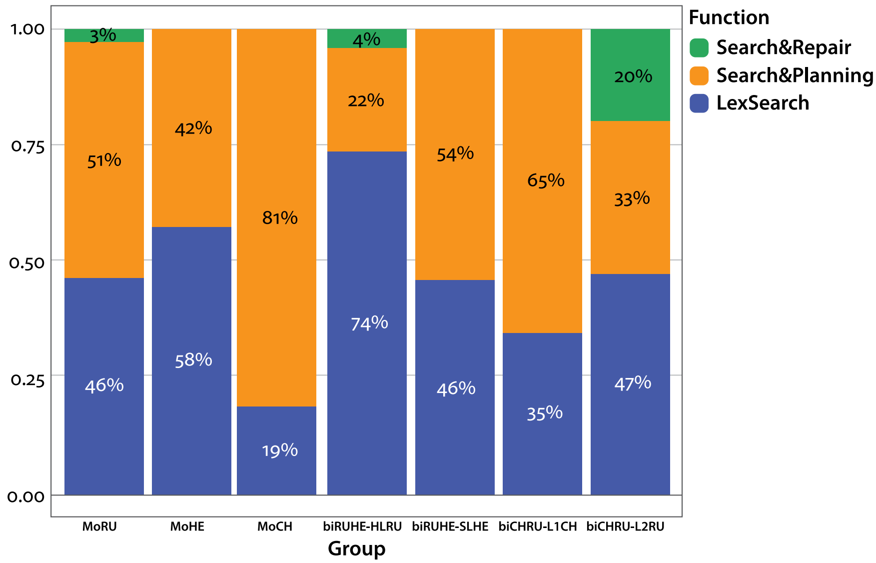

Figure 3. The visualization for the distribution of filled pause function per group
The stacked bar chart shows the relative frequency of each functional category of filled pauses within groups. Values are expressed as percentages of total filled pause per group. Functions are color-coded, and percentages are displayed within each bar segment.

the functional categorization of filled pauses, we did not elaborate extensively on utterance position patterns in the main text. Notably, the distribution of filled pause functions was consistent with broader trends in utterance position. A detailed breakdown of utterance position by group and function is provided in the supplementary materials. Due to data sparsity, contrasts involving utterance-final positions were not estimable for most groups.

3.4 Prolongation position in word and utterance

A linear model examining WordPosition, Group, and their interaction explained a significant proportion of variance in prolongation rates per token ($R^2 = .66$, adj. $R^2 = .59$, $p < .001$; see Table 6 and Figure 4). Word-final prolongations in monolingual Russian speakers (MoRU) served as the baseline (rate = 0.57). Prolongation rates were significantly lower in medial and initial word positions compared to final. Monolingual Hebrew (MoHE), monolingual Chinese (MoCH), and bilinguals producing Mandarin Chinese as L1 (biCHRU-L1CH) showed higher word-final prolongation rates than MoRU, while bilinguals producing Russian as L2 (biCHRU-L2RU) did not differ from MoRU. Post-hoc comparisons revealed MoRU produced fewer word-final prolongations than MoCH and biCHRU-

L1CH; MoHE and MoCH produced more than biCHRU-L2RU; and bilinguals with heritage Russian and societal Hebrew (biRUHE-HLRU, biRUHE-SLHE) produced more than biCHRU-L2RU. These results suggest that word-final prolongation rates in adults are shaped by language dominance and bilingual experience, with distinct patterns depending on language pairings and acquisition history.

Table 6. The output model for the analyses of prolongation position in word per group

Predictor	Estimate	SE	t-value	p-value
(Intercept)	0.57	0.08	7.58	**<0.001**
WordPosition [Medial]	−0.33	0.11	−2.92	**0.004**
WordPosition [Initial]	−0.43	0.14	−2.99	**0.003**
WordPosition [Vowel-like]	−0.12	0.11	−1.11	0.270
Group [MoHE]	0.30	0.10	2.88	**0.005**
Group [MoCH]	0.43	0.11	3.75	**<0.001**
Group [biRUHE-HLRU]	0.15	0.10	1.47	0.145
Group [biRUHE-SLHE]	0.14	0.10	1.40	0.165
Group [biCHRU-L1CH]	0.43	0.10	4.27	**<0.001**
Group [biCHRU-L2RU]	−0.17	0.10	−1.62	0.109
WordPosition [Medial] × Group [MoHE]	−0.42	0.18	−2.33	**0.022**
WordPosition [Vowel-like] × Group [MoHE]	−0.38	0.16	−2.36	**0.020**

Model fit statistics: Multiple $R^2 = 0.66$; Adjusted $R^2 = 0.60$
Note. Only significant results are presented in the table. Significance codes:
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

The visual data in Figure 4 reveal a general cross-linguistic tendency to produce prolongations in word-final position. This pattern likely reflects a universal prosodic strategy, as many final-position words were of the CV type, with vowels that are naturally amenable to prolongation.

Prolongations in medial word position were predominantly observed in the Russian data, particularly on the discourse marker 'vot', which appears to function as a hesitation device akin to 'okay' in English, enabling speakers to maintain the conversational floor while planning upcoming speech. Although most statistical results for non-final positions were not significant, this usage pattern is noteworthy. Among L1-Mandarin Chinese/L2-Russian bilinguals, the elevated rate of vowel-like prolongations in L2-Russian was primarily due to frequent prolongation of the word /i/ 'and.' In contrast, no such pattern was observed in their L1 Mandarin.

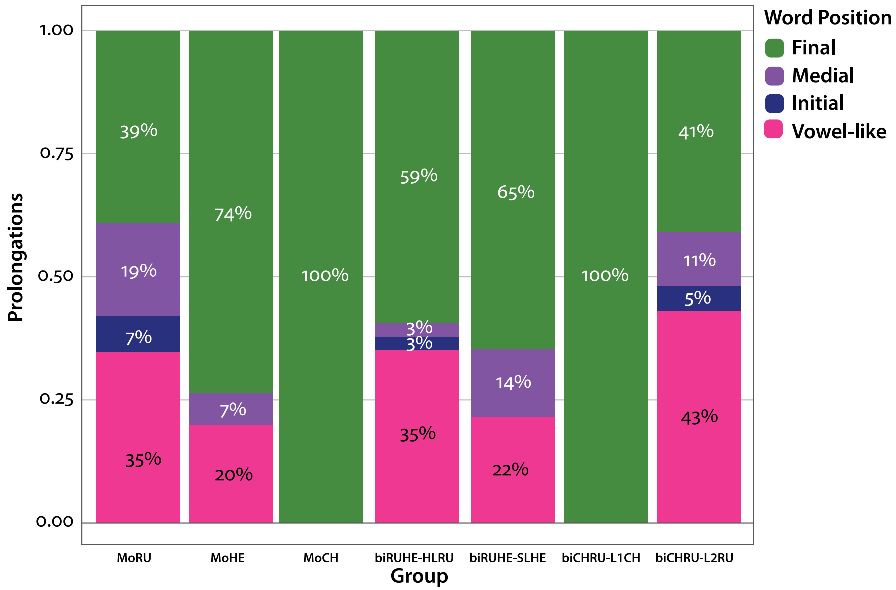

Figure 4. Distribution of prolongations by word position across groups
Proportional distribution of prolongations by word position – initial, medial, final, and vowel-like (i.e., monosyllabic vowel-only words) – across speaker groups. Each bar represents 100% of the prolongations produced by a given group, with segments indicating the relative frequency of each word position category. Percentages are displayed within the bars for clarity.

A linear model (Model B) including UtterancePosition, Group, and their interaction (see Table B and Figure B in the Supplementary Materials) accounted for a significant proportion of the variance in prolongation rates per token ($R^2=.583$, Adjusted $R^2=.501$), with a significant overall effect ($F(20, 102)=7.13$, $p<.001$). The residual standard error was 0.217, indicating a relatively tight fit between the predicted and observed values. In utterance-final position, monolingual Mandarin Chinese (MoCH) speakers produced significantly more prolongations than monolingual Russian (MoRU) speakers ($p=.017$), monolingual Hebrew (MoHE) speakers ($p=.012$), and multiple bilingual groups, including biRUHE-HLRU ($p=.023$), biCHRU-L1CH ($p=.031$), and biCHRU-L2RU ($p=.003$). In utterance-medial position, the biRUHE-HLRU group produced significantly more prolongations than biCHRU-L2RU ($p=.011$), and biCHRU-L1CH had significantly higher rates than biCHRU-L2RU ($p=.050$). These results suggest that Mandarin Chinese speakers – and to some extent Hebrew–Russian bilinguals – demonstrated distinctive disfluency positioning patterns, particularly favoring word-final prolongations, in contrast to monolingual Russian speakers and L2-Russian bilinguals, whose distributions were more balanced across utterance positions.

4. Discussion

Comparisons between monolinguals speaking typologically different languages (Russian, Hebrew, and Mandarin Chinese), as well as between bilinguals and monolinguals, and among bilinguals from different language pairs, are inherently complex due to the multifaceted influences on hesitation usage. For monolinguals, previous research suggested that linguistic diversity contributes to intrinsic variations in hesitation patterns, reflecting differences in language structures and phonotactic rules unique to each language. In bilingual individuals, both internal and external factors play significant roles, including the age of onset of language exposure, cognitive abilities, and language exposure, all of which can impact hesitation usage. Additionally, bilingual speakers may exhibit discrepancies in hesitation patterns between their two languages due to differences in input quantity, input type, and relative proficiency levels (Brundage & Rowe, 2018). In this study, we aimed to explore these comparisons while ensuring similarity in tasks and maintaining homogeneity in age and gender across participant groups.

4.1 Filled pauses and prolongations in monolingual speakers

The results from our first research question, which investigated whether there are differences in hesitation patterns among monolingual speakers of languages with distinct typologies, revealed similar frequencies of filled pauses and prolongations among Russian, Hebrew, and Mandarin Chinese speakers, contrary to our initial prediction. This suggests that the frequencies of filled pauses and prolongations may reflect language-universal properties, potentially rooted in common cognitive processes underlying speech production across languages. Our findings partially align with de Leeuw's (2007) study, which found no significant differences between English and Dutch speakers, and between English and German speakers. However, our results contrast with Böttcher and Zellers (2024), who found that monolingual Russian speakers exhibited the highest rate of filled pauses per 100 words, followed by German and English speakers.

These discrepancies highlight the complexity of identifying universal hesitation patterns, as various factors, including methodological differences, task types, and participant characteristics, can influence the results. In our study, we controlled for these variables by having all participants produce a monologue based on wordless comic strips. Specifically, the consistency of our findings across typologically distinct languages suggests that certain patterns of hesitation – such as the presence of filled pauses and prolongations during cognitively demanding moments to hold the conversational floor – may be universal. These aspects appear to be driven by shared cognitive mechanisms in speech planning and processing. For instance, the narrative task elicited similar hesitation frequencies across age- and gender-balanced groups, except among the L1-Mandarin Chinese/L2-Russian speakers, who faced the greatest cognitive load in their L2-Russian. This group was narrating in a less proficient L2 and had to navigate not only linguistic retrieval but also the cognitive demands of producing a humorous narrative – a genre requiring creativity rather than straightforward description.

Furthermore, 'eh' emerged as the most commonly used filled pause not only in Hebrew, but in all three monolingual groups, including Mandarin Chinese, albeit to a lesser extent. While some types of filled pauses are shown to be language-specific – such as 'en' in Mandarin Chinese – our study suggests that 'eh' is the most common filled pause across monolingual and bilingual speakers of typologically diverse languages, though the degree of preference varies by language.

Regarding the functions of filled pauses, our study found that in all monolingual groups, participants used filled pauses for both *Lexical Search* and *Search and Planning*. These functions were consistent across Russian, Hebrew, and Mandarin Chinese speakers, indicating that the role of filled pauses in facilitating cognitive processes during speech is a language-universal characteristic. Filled pauses help speakers manage the complexities of speech production, such as retrieving words and planning utterances. While prior research (e.g., Clark & Fox Tree, 2002; Shriberg, 2001) has established these functions primarily in English and other individual languages, our study is the first to systematically compare filled pause functions across typologically diverse languages and across both monolingual and bilingual adults using the same task. This unified approach allows for a controlled cross-linguistic and cross-population comparison, providing new empirical support for the universal nature of filled pause functions in speech production. While the specific types of filled pauses may vary across languages, their underlying communicative and cognitive roles appear to be universally shared.

We anticipated that Mandarin Chinese speakers would avoid bilabial nasal-vocalic filled pauses due to phonotactic constraints that prohibit word-final [m] in native vocabulary (Wu & Kenstowicz, 2015). Interestingly, nasal-vocalic filled pauses were almost entirely absent among monolingual Mandarin Chinese speakers and occurred only rarely in L1-Mandarin Chinese/L2-Russian bilinguals. As predicted, we observed a high frequency of word-final prolongations among both monolingual and bilingual speakers of Hebrew and Mandarin Chinese (Lee et al., 2004; Silber-Varod et al., 2019), in contrast to Russian speakers, whose prolongation distribution was more evenly spread across word positions (Bogdanova-Beglarian et al., 2023).

4.2 Filled pauses and prolongations in bilingual speakers as compared to monolingual controls

Our second research question examined whether there are differences in the rates of filled pauses and prolongations between bilinguals and monolinguals in their respective languages. As anticipated, bilingual L1-Mandarin Chinese/L2-Russian speaking bilingual speakers exhibited a higher frequency of filled pauses and prolongations in their L2-Russian compared to the monolingual Russian controls, while no increase in hesitations was observed in their stronger L1-Mandarin Chinese relative to monolingual Mandarin Chinese speakers. However, the findings for HL-Russian/SL-Hebrew bilinguals did not fully align with our expectations. HL-Russian/SL-Hebrew speakers did not differ significantly from monolingual Hebrew-speaking controls in their use of either filled pauses or prolongations in SL-Hebrew, the stronger language. In contrast, although we had expected HL-Russian speakers to exhibit more hesitations than monolingual Russian speakers – given that HL-Russian is typically the weaker language – we found no significant differences between the groups in their use of either filled pauses or prolongations. These findings contrast with those of Böttcher and Zellers (2024), who reported higher frequencies of filled pauses in the narratives of bilingual heritage speakers compared to monolinguals in the same language. The findings for the two bilingual groups align with the Critical Period hypothesis, which posits monolingual-like performance for early bilinguals (like in this study for HL-Russian/ SL-Hebrew bilinguals), whereas divergence for late bilinguals in their L2 (like for L1-Mandarin Chinese/L2-Russian bilinguals).

Notably, among L1-Mandarin Chinese/L2-Russian speakers, proficiency in L2-Russian was significantly weaker compared to their L1-Mandarin Chinese and to monolingual Russian-speaking controls. Meanwhile, among HL-Russian/SL-Hebrew speakers, proficiency in HL-Russian was relatively weaker compared to SL-Hebrew and monolingual Hebrew-speaking controls. These findings are consistent with and build upon existing bilingual speech production models (e.g., Kroll & Bialystok, 2013), which propose that bilinguals face increased cognitive demands that may affect hesitation patterns both quantitatively and qualitatively. For example, the variation in forms of filled pauses and the placement of prolongations may reflect differential access to phonological representations or cross-linguistic influence, illustrating bilinguals' dynamic speech planning processes.

The overall tendency across all groups to prolong segments in final word position, particularly in CV-type or vowel-final words, suggests a language-universal prosodic strategy. Such prolongations may serve to maintain speech rhythm, signal continuation, or allow planning time, aligning with cross-linguistic evidence that prosodic boundary regions are cognitively demanding and thus prone to hesitation phenomena.

At the same time, several patterns point to language-specific influences. For example, the L2-Russian showed a distinct preference for prolonging sonorant consonants and glides, diverging from the more vowel-centered prolongations in Hebrew and Russian. This may reflect phonotactic constraints or articulatory patterns more common in Mandarin, as well as L2-specific processing effects. Furthermore, medial-position prolongations, while rare overall, were primarily found in Russian, particularly on the discourse marker *vot*. This item appears to have acquired a language-specific discourse function analogous to hesitation markers like *okay* in English, supporting the idea that certain hesitation strategies are shaped by language-internal pragmatics and lexical routines. Furthermore, the differences in hesitation patterns between heritage speakers and those learning a second language later in life highlight the importance of early and sustained language exposure in achieving fluency and effectively managing hesitations. The contrasting results between our study and others, such as Böttcher and Zellers (2024), further illustrate the complex interplay of linguistic, cognitive, and environmental factors that shape hesitation patterns among bilingual and monolingual speakers.

4.3 Filled pauses and prolongations in the two languages of bilingual speakers

The third research question examined whether differences in filled pauses and prolongations exist between the two languages of bilinguals, comparing HL-Russian and SL-Hebrew in heritage bilinguals, as well as L1-Mandarin Chinese and L2-Russian in Mandarin Chinese-Russian bilinguals. Significant differences were found only in hesitation frequency between L1-Mandarin Chinese and L2-Russian, suggesting that cognitive load and planning difficulties impact the weaker L2 (Russian) more than the stronger L1 (Mandarin Chinese). This finding aligns with previous studies, such as Muhlack's (2023) investigation of Spanish-English bilinguals, which reported a higher frequency of filled pauses in the speakers' L2. Similarly, Teng and Androsova (2022) demonstrated that while both Mandarin Chinese and Russian speakers exhibited similar ratios of silent pauses to filled pauses in their L1, this ratio shifted significantly when Mandarin Chinese speakers switched to L2-Russian.

No significant differences were found between HL-Russian and SL-Hebrew in heritage bilinguals, which may be attributed to their early and simultaneous exposure to both languages, as well as their relatively high proficiency in HL-Russian. These findings align with studies involving advanced adult learners, such as De Boer and Heeren (2020), who examined native Dutch speakers with high proficiency in L2-English, as well as Peltonen et al. (2023), who studied advanced adult learners of L2-English with L1-Finnish, and Cuartero et al. (2023), who focused on heritage Spanish speakers in both HL-Spanish and SL-English. Together, these results highlight the impact of both proficiency (as seen in De Boer and Heeren, 2020, and Peltonen et al., 2023) and early exposure to both languages (as in Cuartero et al., 2023) on disfluency production patterns. However, our findings contrast with Guz (2015), who observed a higher rate of filled pauses in L2 speech, despite the participants' advanced English proficiency. This discrepancy may stem from differences in L2 exposure: while De Boer and Heeren (2020) studied students immersed in a multilingual campus where English was the lingua franca, Guz's Polish students experienced L2-English primarily through structured classroom learning.

Filled pause form analysis confirmed monolingual-like use of filled pauses and prolongations in HL speakers. For example, HL-Russian/SL-Hebrew speakers did not produce the filled pause 'a' in SL-Hebrew, as it is not used by Hebrew monolinguals, but they did produce 'a' in their HL-Russian, similar to Russian monolinguals. In L1-Mandarin Chinese/L2-Russian speakers, as expected, the 'em' variant was virtually absent among both L1-Mandarin Chinese speakers and monolingual Mandarin speakers, who instead tend to use 'en', consistent with the phonotactic constraints of Mandarin (Yuan et al., 2016). Moreover, in L2 speakers, a transfer effect was observed in L1-Mandarin Chinese/L2-Russian bilinguals in both directions. The Mandarin-specific form 'en' was transferred into their L2-Russian. However, several instances of 'am' ('a-am') and 'm' ('m-m') were observed in L2-Russian, where such filled pause forms are phonotactically permissible, but not in Mandarin Chinese. Interestingly, filled pause types analyses revealed that L1-Mandarin Chinese/L2-Russian bilinguals used filled pauses not observed in either monolingual Russian or Mandarin Chinese controls, such as 'y' ('y-y') in L1 and 'n-n' and 'an' ('a-an') in L2-Russian. These hybrid variants suggest blending Russian 'a-a' and 'em' into 'a-an', and converting Russian 'm-m' to 'n-n'.

Additionally, filled pause function analysis revealed that Russian-Hebrew heritage speakers produced filled pauses more often for *Lexical Search* in HL-Russian. These findings reveal that Russian-Hebrew speakers face specific challenges in lexical search, while discourse and syntactic planning appear comparatively less demanding. In contrast, Mandarin Chinese-Russian speakers encounter difficulties primarily in speech planning, requiring more time to organize entire utterances. Notably, our results align with Zhao and Jurafsky's (2005) findings, which show that Mandarin Chinese speakers frequently use demonstratives in nominal-searching contexts, while filled pauses are more commonly found clause-initially. Together, these insights underscore the influence of linguistic and cognitive factors on hesitation patterns, highlighting how language-specific structures shape the interplay between lexical retrieval and broader speech planning processes.

5. Limitations

While our study utilized a unified methodology across all participant groups to compare filled pause and prolongation rates, it is important to acknowledge that variations in filled pause patterns may still arise due to additional factors. For instance, Mandarin Chinese-Russian students may experience heightened stress during evaluations by academic authorities, potentially influencing their filled pause patterns compared to HL speakers who may feel more comfortable in their native environment. Moreover, a comparison of bilinguals with different L1 backgrounds but the same L2, such as native Hebrew speakers paired with various L2s, could offer valuable insights. Additionally, comparing HL/SL and L1/L2 speakers of different language pairs building on Götz's (2019) findings, which highlighted the impact of different L1 backgrounds on filled pause production in L2 among advanced English learners – could reveal distinct patterns across diverse groups. A further limitation of our study is the exclusion of silent pauses, which, when analyzed alongside filled pauses, could offer a more nuanced understanding of differences in speech production and planning processes. Future research integrating silent pauses could provide a richer picture of hesitation phenomena across languages and contexts.

6. Conclusions

The present study explored hesitation patterns among monolingual and bilingual speakers from different language backgrounds. Our findings highlight the complexity of hesitation usage. The three monolingual groups showed similar rates of filled pauses and prolongations, despite differences in hesitation patterns reported in previous studies, suggesting these phenomena may reflect language-universal features of hesitation. However, the form and function of filled pauses, as well as the position of prolongations within words and utterances, showed tendencies to vary by language, indicating language-specific patterns.

Among bilingual speakers, our results underscore the significance of internal and external factors — such as the age of bilingualism onset, language proficiency, and exposure — in shaping hesitation patterns. No differences between HL-Russian and monolingual Russian in the frequency of filled pauses and prolongations reflect the critical role of early and sustained exposure to HL in the production of speech hesitations, in line with the Critical Period Hypothesis. The study also revealed that bilinguals exhibit varied hesitation patterns across their two languages, particularly when one language is significantly weaker. This was evident in the higher hesitation rates observed in L2-Russian compared to L1-Mandarin Chinese among L1-Mandarin Chinese/L2-Russian bilinguals in accordance with research that shows that bilingual adults tend to produce more disfluencies when speaking in their non-native language (e.g., Bergmann et al., 2015; Gkalitsiou & Werle, 2023; Smith et al., 2022), particularly in more demanding tasks and contexts. In addition, the findings emphasize the adaptive nature of hesitation usage in bilingual speech, shaped by linguistic proficiency, phonotactic constraints, and exposure to the target language.

Overall, this study contributes to the understanding of hesitation patterns by demonstrating the interplay between universal cognitive processes and language-specific influences.

Acknowledgments

This research was supported by the Fundamental Research Fund for the Central Universities of China (Grant No. 2242025S30052, Xiaoli Sun), by Saint Petersburg State University (Project No. 124032900006-1, Modeling of Russian Megalopolis Citizens' Communicative Behavior in Social, Speech and Pragmatic Aspects Using Artificial Intelligence Methods, N. V. Bogdanova-Beglarian), and was partially supported by the Israel Science Foundation (ISF; Grant No. 552/21, Towards Understanding Heritage Language Development: The Case of Child and Adult Heritage Russian in Israel and the USA, Natalia Meir).

References

Akhutina, T.V. (2007). Vygotsky-Luria-Leontiev's school of psycholinguistics: The mechanisms of language production. In R. Alanen & S. Pöyhönen (Eds.), *Language in action: Vygotsky and Leontievian legacy today* (pp. 32–56). Cambridge Scholars Publishing.

doi Anderson, A.H., Bader, M., Bard, E.G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H.S., & Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech*, *34*(4), 351–366.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*, *32*(1), 25–36.

Asherov, D., & Bat-El, O. (2019). Syllable structure and complex onsets in Modern Hebrew. *Journal of Afroasiatic Languages and Linguistics*, *11*(1), 69–95.

Aswini, V., Selvaraj, J. L., Prathiba, P., & Reji, A. (2022). Disfluencies and rate characteristics of English speaking bilingual young Indian adults: Effects of task and speech situations. *International Journal of Science & Healthcare Research*, *5*(2), 128–137.

Azogui, J., Lerner, A., & Silber-Varod, V. (2015). *The Open University of Israel Map Task Corpus (MaTaCOp)*. The Open University of Israel. http://www.openu.ac.il/en/academicstudies/matacop/

Beradze, M., & Meir, N. (2025). Supplementary materials for "An attempt to identify language-universal and language-specific patterns in the use of filled pauses and prolongations" [Data set, R script, and analysis output]. *Open Science Framework*. https://osf.io/afrc8/?view_only=63bc6503b0754fadbacf246150cc9630

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bedore, L. M., Fiestas, C. E., Peña, E. D., & Nagy, V. J. (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition*, *9*(3), 233–247.

Bergmann, C., Sprenger, S. A., & Schmid, M. S. (2015). The impact of language co-activation on L1 and L2 speech fluency. *Acta Psychologica*, *161*, 25–35.

Betz, S., Bryhadyr, N., Kosmala, L., & Schettino, L. (2021). A crosslinguistic study on the interplay of fillers and silences. In R.L. Rose & R. Eklund (Eds.), *Proceedings of DiSS 2021: The 10th Workshop on Disfluency in Spontaneous Speech* (pp. 47–52). Université Paris 8 Vincennes-Saint-Denis.

Betz, S., Bryhadyr, N., Türk, O., & Wagner, P. (2023). Cognitive Load Increases Spoken and Gestural Hesitation Frequency. *Languages*, *8*(1), Article 71.

Betz, S., Eklund, R., & Wagner, P. (2017). Prolongation in German. In R. Eklund & R. Rose (Eds.), *TMH-QPSR: Vol. 58. Proceedings of DiSS 2017, Disfluency in Spontaneous Speech* (pp. 13–16). Royal Institute of Technology, Sweden.

Betz, S., & Wagner, P. (2016). Disfluent lengthening in spontaneous speech. In *Elektronische Sprachsignalverarbeitung (ESSV) 2016*.

Bick, A. S., Goelman, G., & Frost, R. (2011). Hebrew brain vs. English brain: language modulates the way it is processed. *Journal of cognitive neuroscience*, *23*(9), 2280–2290.

Bidstrup, H. (2008). *Izbrannoe* [*Selected works*]. Moskva: Izdatel'skii Dom Meshcheryakova.

Bihovsky, A., Ben-Shachar, M., & Meir, N. (2025). Within- and cross-language generalization in narrative production of bilingual persons with aphasia following semantic feature analysis therapy. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *77*(3), 284–299.

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version unknown. Retrieved from http://www.praat.org/

Bogdanova-Beglarian, N. V., Brodt, I.S., Kukanova, V.V., Pavlova, O.V., Sapunova, E.M., Subbota, I.A., Filippova, N.S., Khan, N.A., & Chujko, V.M. (Eds.). (2013). *Zvukovoj korpus kak material dlya analiza russkoj rechi. Chast' 1. Chtenie. Pereskaz. Opisanie* [*The sound corpus as material for the analysis of Russian speech. Part 1: Reading. Retelling. Description*]. Saint Petersburg: Saint Petersburg State University.

Bogdanova-Beglarian, N., & Baeva, E. (2018). Nonverbal elements in everyday Russian speech: An attempt at categorization. In A. Elizarov & N. Loukachevitch (Eds.), *Computational models in language and speech: Proceedings of Computational Models in Language and Speech Workshop (CMLS 2018)* (pp. 3–13). Kazan, Russia: Kazan (Volga Region) Federal University.

Bogdanova-Beglarian, N., Zaides, K., Stoika, D., & Sun, X. (2023). Prolongations as hesitation phenomena in spoken speech in first and second language. In A. Karpov et al. (Eds.), *Speech and Computer. SPECOM 2023. Lecture Notes in Computer Science* (Vol. *14338*, pp. 357–367). Springer.

Bogdanova-Beglarian, N., Zaides, K., Verkhovtceva, T., Beradze, M., & Meir, N. (2022). Self-repair in elicited narrative production in speakers of Russian as the first (L1), second (L2), and heritage (HL) language. *Languages*, *7*(3), Article 229.

Bortfeld, H., Leon, S.D., Bloom, J.E., Schober, M.F., & Brennan, S.E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, *44*(2), 123–147.

Böttcher, M., & Zellers, M. (2024). Do you say uh or uhm? A cross-linguistic approach to filler particle use in heritage and majority speakers across three languages. *Frontiers in Psychology*, *15*, 1305862.

Braver, A., Dresher, N., & Kawahara, S. (2016). The Phonetics of Emphatic Vowel Lengthening in English. *Proceedings of the Annual Meetings on Phonology*, *2*.

Brundage, S.B., & Rowe, H. (2018). Rates of typical disfluency in the conversational speech of 30-month-old Spanish-English simultaneous bilinguals. *American Journal of Speech-Language Pathology*, *27*(3S), 1287–1298.

Chen, C. (2017). *Hezitatsii v russkoi ustnoi rechi nositelei kitaiskogo yazyka* [*Hesitations in Russian spoken speech of Mandarin Chinese speakers*]. Unpublished doctoral dissertation, Saint Petersburg State University, Saint Petersburg. Retrieved from https://disser.spbu.ru/files/disser2/disser/85or4wMfLM.pdf

Clark, H. H. (2013). Conceptions of disfluencies [Plenary abstract]. In R. Eklund (Ed.), *Proceedings of DiSS 2013: The 6th Workshop on Disfluency in Spontaneous Speech (TMH-QPSR, 54[1])*. KTH Royal Institute of Technology. http://roberteklund.info/conferences/diss2013

Clark, H.H., & Fox Tree, J.E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, *84*(1), 73–111.

Clark, H.H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, *37*(3), 201–242.

Cuartero, M., Domínguez, M., & Pascual y Cabo, D. (2023). Examining Oral (Dis)Fluency in — uh- — Spanish as a Heritage Language. *Languages*, *8*(3), 173.

De Boer, M.M., & Heeren, W. (2020). Cross-linguistic filled pause realization: The acoustics of uh and um in native Dutch and non-native English. *The Journal of the Acoustical Society of America*, *148*(6), 3612–3622.

De Leeuw, E. (2007). Hesitation Markers in English, German, and Dutch. *Journal of Germanic Linguistics*, *19*(2), 85–114.

Deme, A. (2013). Pragmatic functions of lengthenings and filled pauses in the adult-directed speech of Hungarian children. In *New Perspectives on Speech in Action: 2nd SJUSK Conference on Contemporary Speech Habits* (pp. 23–40). Samfundslitteratur.

Didirková, I. (2024). Disfluency in speech and language disorders. *Clinical Linguistics & Phonetics*, *38*(4), 285–286.

Eklund, R. (2001). Prolongations: A dark horse in the disfluency stable. In *Proceedings of ITRW on Disfluency in Spontaneous Speech (DiSS 2001)* (pp. 5–8).

Eklund, R. (2004). *Disfluency in Swedish human-human and human-machine travel booking dialogues* [Doctoral dissertation, Linköping University]. Linköping University Electronic Press (DiVA portal). https://liu.diva-portal.org/smash/get/diva2:20923/FULLTEXT01.pdf

Fehringer, C., & Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica*, *41*(1–2), 37–72.

Felker, E. R., Klockmann, H. E., & De Jong, N. H. (2019). How conceptualizing influences fluency in first and second language speech production. *Applied Psycholinguistics*, *40* (1), 111–136.

Fiestas, C. E., Bedore, L. M., Peña, E. D., & Nagy, V. J. (2005). Use of mazes in the narrative language samples of bilingual and monolingual 4- to 7-year old children. In J. Cohen et al. (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 730–740). Cascadilla Press.

Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, *29*(3), 320–326.

Fox Tree, J. E. (2006). Disfluencies in spoken language. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. John Wiley & Sons, Ltd.

Fraser, B. (2006). Towards a theory of discourse markers. In *K. Fischer Approaches to discourse particles* (pp. 189–204). Brill.

García-Amaya, L., & Lang, S. (2020). Filled pauses are susceptible to cross-language phonetic influence: Evidence from Afrikaans-Spanish bilinguals. *Studies in Second Language Acquisition*, *42*(5), 1077–1105.

Gilquin, G. (2008). Hesitation markers among EFL learners: Pragmatic deficiency or difference?. In J. Romero-Trillo (Ed.), *Pragmatics and Corpus Linguistics: A mutualistic entente* (pp. 119-150). Berlin, New York: De Gruyter Mouton.

Gkalitsiou, Z., & Werle, D. (2023). Speech disfluencies in bilingual Greek-English young adults. *Journal of Fluency Disorders*, *78*, Article 106001.

Gósy, M. (2023). Occurrences and durations of filled pauses in relation to words and silent pauses in spontaneous speech. *Languages*, *8*(1), Article 79.

Gósy, M., & Eklund, R. (2017). Segment prolongation in Hungarian. In R. Eklund & R. L. Rose (Eds.), *Proceedings of DiSS 2017: The 8th Workshop on Disfluency in Spontaneous Speech* (TMH-QPSR, Vol. 58(1), pp. 29–32). KTH Royal Institute of Technology.

Götz, S. (2019). Do learning context variables have an effect on learners' (dis)fluency? Language-specific vs. universal patterns in advanced learners' use of filled pauses. In E. Degand, L. Meurant, G. Gilquin, & A. C. Simon (Eds.), *Fluency and Disfluency across Languages and Language Varieties* (pp. 177–196). Presses universitaires de Louvain.

Gulimire, A., & Titkova, S. I. (2018). The Russian language in China: Traditions and prospects. *Modern Linguistic and Methodical-and-Didactic Researches*, *2*(38), 68–75. Retrieved from https://elibrary.ru/item.asp?id=35257725

doi Guz, E. (2015). Establishing the fluency gap between native and non-native speech. *Research in Language*, *13*(2), 230–247.

doi Hernandez Santacruz, J., Torres, J., & Kroll, J. F. (2025). Heritage speakers reveal the dynamics of bilingual language regulation. *Bilingualism: Language and Cognition*, 1–14.

doi Hilton, M., Räling, R., Wartenburger, I., & Elsner, B. (2019). Parallels in Processing Boundary Cues in Speech and Action. *Frontiers in Psychology*, *10*, 458615.

doi Hlavac, J. (2011). Hesitation and monitoring phenomena in bilingual speech: A consequence of code-switching or a strategy to facilitate its incorporation? *Journal of Pragmatics*, *43*(15), 3793–3806.

Holden, K. T. (1978). Initial and final consonant clusters in Russian and English. *Russian Language Journal*, *32*(112), 19–42. http://www.jstor.org/stable/43909390

doi Johnson, J. S., & Newport, E. L. (1988). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99.

doi Kentner, G., Franz, I., Knoop, C. A., & Menninghaus, W. (2023). The final lengthening of pre-boundary syllables turns into final short-ening as boundary strength levels increase. *Journal of Phonetics*, *97*, 101225.

Kibrik, A. A., & Podlesskaja, V. I. (Eds.). (2009). *Rasskazy o snovidenijax: Korpusnoe issledovanie ustnogo russkogo diskursa* [*Night dream stories: A corpus study of spoken Russian discourse*]. Moscow: Jazyki slavjanskix kul'tur. Retrieved from http://www.lrc-press.ru/pics/previews/ru/(165)Pages%20from%20Sb-Rasskazy-2009-v.pdf

doi Kogan, N., & Saiegh-Haddad, E. (2023). The internal structure of the syllable in Russian and in Hebrew: Evidence from monolingual kindergarteners. *Applied Psycholinguistics*, *44*(1), 101–129.

Kosmala, L., & Morgenstern, A. (2019). Should uh and um be categorized as markers of disfluency? The use of fillers in a challenging conversational context. In L. Degand, G. Gilquin, L. Meurant, & A. C. Simon (Eds.), *Fluency and Disfluency across Languages and Language Varieties* (pp. 67–90). Presses universitaires de Louvain.

doi Kosmala, L., & Crible, L. (2022). The dual status of filled pauses: Evidence from genre, proficiency, and co-occurrence. *Language and Speech*, *65*(1), 216–239.

doi Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514.

doi Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. H. Ross (Ed.), The psychology of learning and motivation (Vol. *56*, pp. 229–262). Academic Press.

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick & V. S. Ferreira (Eds.), *The Oxford Handbook of Language Production* (pp. 165–181). Oxford University Press.

Lee, T.-L., He, Y.-F., Huang, Y.-J., Tseng, S.-C., & Eklund, R. (2004). Prolongation in spontaneous Mandarin. In *Proceedings of Interspeech 2004* (pp. 2181–2184).

doi Lenneberg, E. H. (1967) *Biological foundations of Language*. New York: John Wiley & Sons.

Lickley, R.J. (2015). Fluency and disfluency. In M.A. Redford (Ed.), *The Handbook of Speech Production* (pp. 445–474). John Wiley & Sons, Inc.

Llisterri, J., Machuca, M.J., & Ríos, A. (2022). La función de las hesitaciones en la identificación del hablante. In B. Blecua, J. Cicres, M. Espejel, & M.J. Machuca (Eds.), *Propuestas en fonética experimental: enfoques metodológicos y nuevas tecnologías* (pp. 160–164). Universitat de Girona. http://hdl.handle.net/10256/20770

Lo, J.J.H. (2020). Between Äh(m) and Euh(m): The distribution and realization of filled pauses in the speech of German–French simultaneous bilinguals. *Language and Speech*, *63*(4), 746–768.

Maclay, H., & Osgood, C.E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, *15*(1), 19–44.

Maekawa, K. (2004). Outline of the Corpus of Spontaneous Japanese. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous Speech: Data and Analysis* (pp. 11–15). National Institute for Japanese Language.

Montrul, S., & Polinsky, M. (2021). Introduction — heritage languages, heritage speakers, heritage linguistics. In S. Montrul & M. Polinsky (Eds.), *The Cambridge Handbook of Heritage Languages and Linguistics* (pp. 1–10). Cambridge University Press.

Muhlack, B. (2023). Filler particles in English and Spanish L1 and L2 speech. In *Proceedings of ICPhS 2023* (pp. 2418–2422). GUARANT International.

Muhlack, B., Trouvain, J., & Jessen, M. (2023). Distributional and acoustic characteristics of filler particles in German with consideration of forensic-phonetic aspects. *Languages*, *8*(2), Article 100.

Newport, E.L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28.

Peck, N. & Becker, L. (2024). Syntactic pausing? Re-examining the associations. *Linguistics Vanguard*, *10*(1), 223–237.

Peltonen, P., Kosmala, L., Götz, S., & Lintunen, P. (2023). The interplay between speech fluency and gesture in L1 Finnish and L2 English task-based interactions. In *Proceedings of DiSS 2023* (pp. 8–12).

Pletikosa, J., & Rungrojsuwan, S. (2018). The use of filled pauses in monologue by Thai learners of English with different English language proficiency levels. *Vacana Journal of Language and Linguistics*, *6*(25), 20–35.

Podlesskaja, V.I., & Kibrik, A.A. (2007). Samoispravleniya govoryashchego i drugie tipy rechevykh sboev kak ob"ekt annotirovaniya v korpusakh ustnoy rechi [Self-corrections of a speaker and other types of disfluencies in the corpus of oral speech]. *Nauchno-tekhnicheskaya informatsiya. Seriya 2: Informatsionnye protsessy i sistemy* [*Scientific and Technical Information. Series 2: Information Processes and Systems*], (2), 2–23.

Qi, N. (2021). The interaction of foreign languages and society: A review of Russian language education in China. *Onomázein*, Special Issue IX — Foreign Languages Education in China: Current Situation and Future Prospects, 20–37.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, *13*(2), 155–163.

Saint Petersburg State University. (n.d.). Test of Russian as a foreign language (TORFL). Retrieved from https://testingcenter.spbu.ru/en/exams/russian/torfl.html

Savino, M., & Refice, M. (2000). Acoustic cues for classifying communicative intentions in dialogue systems. In *Text, Speech and Dialogue* (pp. 421–426). Springer.

Schettino, L., & Eklund, R. (2023). Prolongation in Italian. In S. Betz, B. Ludusan, & P. Wagner (Eds.), *Proceedings of DiSS 2023: The 11th Workshop on Disfluency in Spontaneous Speech* (pp. 81–85). International Speech Communication Association (ISCA).

Schettino, L., Alfano, I., Cataldo, V., & Leo, G. (2022). A crosslinguistic study on filled pauses and prolongations in Italian and Spanish. In R. Orrico & L. Schettino (Eds.), *La posizione del parlante nell'interazione: atteggiamenti, intenzioni ed emozioni nella comunicazione verbale = The position of the speaker in interaction: attitudes, intentions, and emotions in verbal communication* (Vol. 9, pp. 171–189).

Schnadt, M. J., & Corley, M. (2006). The influence of lexical, conceptual and planning based factors on disfluency production. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *28*(28), 2199–2204. Retrieved from https://escholarship.org/uc/item/9337x2hk

Schüppert, A., & Heisterkamp, P. (2021). Lecturing in L1 Dutch and in L2 English: A pairwise comparison of speech samples of three HE lecturers. *Journal of English for Academic Purposes*, *52*, Article 100993.

Shriberg, E. (2001). To 'errrr' is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, *31*(1), 153–169.

Silber-Varod, V. (2011). The SpeeCHain perspective: Prosodic-syntactic interface in spontaneous spoken Hebrew [Doctoral dissertation, Tel Aviv University]. https://www.openu.ac.il/personal_sites/vered-silber-varod/download/Vered%20Silber-Varod%20Dissertation-7.pdf

Silber-Varod, V. (2013). Structural analysis of prosodic pattern: The case of excessive prolongations in Israeli Hebrew. *Revista Leitura*, *52*, 271–291.

Silber-Varod, V., Gósy, M., & Eklund, R. (2019). Segment prolongation in Hebrew. In R. L. Rose & R. Eklund (Eds.), *Proceedings of DiSS 2019: The 9th Workshop on Disfluency in Spontaneous Speech* (pp. 47–50). Eötvös Loránd University.

Silber-Varod, V., Gósy, M., & Lerner, A. (2021). Is it a filler or a pause? A quantitative analysis of filled pauses in Hebrew. In A. Karpov, K. Samudravijaya, & R. M. Hegde (Eds.), *Proceedings of SPECOM 2021* (pp. 638–648). Springer International Publishing.

Smith, S. A., Choo, A. L., & Seitz, S. R. (2022). Disfluencies in English speech produced by Spanish-English bilinguals. *Folia Phoniatrica et Logopaedica*, *74*(2), 122–130.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of memory and language*, *32*(1), 25–38.

Strassel, S., Kolár, J., Song, Z., Barclay, L., & Glenn, M. (2005). Structural metadata annotation: Moving beyond English. *Proceedings of Interspeech 2005*, 1545–1548.

Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics*, *30*, 485–496.

Teng, H., & Androsova, S. V. (2022). Pauses in native and foreign-accented spontaneous speech: Acoustic analysis of native Chinese, Chinese learners of Russian, and native Russian speech. *Vestnik NSU. Series: History and Philology*, *21*(2), 67–86.

Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders*, *23*(2), 309–322.

Torreira, F., Adda-Decker, M., & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, *52*(3), 201–212.

Tottie, G. (2016). Planning what to say: Uh and um among the pragmatic markers. In G. Kaltenböck, E. Keizer, & A. Lohmann (Eds.), *Outside the Clause: Form and Function of Extra-Clausal Constituents* (pp. 97–122). John Benjamins Publishing Company.

Vasilescu, I., & Adda-Decker, M. (2007). A cross-language study of acoustic and prosodic characteristics of vocalic hesitations. In A. Esposito, M. Bratanić, E. Keller, & M. Marinaro (Eds.), *Fundamentals of verbal and nonverbal communication and the biometric issue (NATO Security through Science Series – E: Human and Societal Dynamics*, Vol. *18*, pp. 140–148). Amsterdam: IOS Press. Retrieved from https://ebooks.iospress.nl/volume/fundamentals-of-verbal-and-nonverbal-communication-and-the-biometric-issue

Verkhodanova, V., Karpov, A., Kipyatkova, I., Shapranov, V. (2018). Automatic detection of vocalized hesitations in Russian speech. *Voprosy Jazykoznanija* (6), pp.104–118.

Vitale, V. N., Schettino, L., & Cutugno, F. (2024). Modelling filled particles and prolongation using end-to-end automatic speech recognition systems: A quantitative and qualitative analysis. In F. Dell'Orletta, A. Lenci, S. Montemagni, & R. Sprugnoli (Eds.), *Proceedings of the 10th Italian Conference on Computational Linguistics (CLiC-it 2024)* (pp. 990–996). CEUR Workshop Proceedings. https://aclanthology.org/2024.clicit-1.107/

Wang, J., Schiller, N. O., & Verdonschot, R. G. (2024). Morphological encoding in language production: Electrophysiological evidence from Mandarin Chinese compound words. *PloS One*, *19*(10), e0310816.

Wieling, M., Grieve, J., Bouma, G., Fruehwald, J., Coleman, J., & Liberman, M. (2016). Variation and change in the use of hesitation markers in Germanic languages. *Language Dynamics and Change*, *6*(2), 199–234.

Wu, F., & Kenstowicz, M. (2015). Duration reflexes of syllable structure in Mandarin. *Lingua*, *164*, 87–99.

Yuan, J., Xu, X., Lai, W., & Liberman, M. (2016). Pauses and pause fillers in Mandarin monologue speech: The effects of sex and proficiency. In *Proceedings of Speech Prosody 2016* (pp. 1167–1170). International Speech Communication Association (ISCA).

Zhao, Y., & Jurafsky, D. (2005). A preliminary study of Mandarin filled pauses. In *Proceedings of DiSS'05: Disfluency in Spontaneous Speech Workshop* (pp. 179–182). Aix-en-Provence, France. https://www.isca-archive.org/diss_2005/zhao05_diss.pdf

Zhao, Z. (2023). Khezitatsionnye yavleniia v russkom ustnom monologe-rasskaze kitaitsev v zerkale sotsio- i psikholingvistiki [Hesitation phenomena in the Russian oral monologue-story of Chinese people through sociolinguistics and psycholinguistics]. *Vestnik BGU. Filologiia*, (4), 36–44.

Address for correspondence

Marianna Beradze
Bar-Ilan University
Israel
beradzem@gmail.com
https://orcid.org/0000-0001-6028-5915

Co-author information

Tatiana Verkhovtceva
Bar Ilan University
tan.vero4@gmail.com
https://orcid.org/0000-0003-1225-8298

Natalia Bogdanova-Beglarian
S-Petersburg State University
nvbogdanova_2005@mail.ru
https://orcid.org/0000-0002-7652-0358

Xiaoli Sun
Southeast University
sunxiaoli_smile@163.com
https://orcid.org/0000-0003-2319-7398

Natalia Meir
Bar Ilan University
natalia.meir@biu.ac.il
https://orcid.org/0000-0001-9426-811X

Kristina Zaides
zaides.kristina@gmail.com
https://orcid.org/0009-0002-4088-910X

Publication history

Date received: 18 November 2024
Date accepted: 18 July 2025
Published online: 6 September 2025