frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 29 August 2024
DOI 10.3389/fcomm.2024.1371126

OPEN ACCESS

EDITED BY
Alicia Luque,
Nebrija University, Spain

REVIEWED BY
Elena Babatsouli,
University of Louisiana at Lafayette, United States
Francesco Romano,
Halmstad University, Sweden

*CORRESPONDENCE
Sara Košutar
✉ sara.kosutar@uit.no

RECEIVED 15 January 2024
ACCEPTED 15 August 2024
PUBLISHED 29 August 2024

CITATION
Hržica G, Košutar S and Poropat Jeletić N (2024) The relationship between self-assessment of language proficiency and measures of lexical diversity and syntactic complexity: evidence from bilingual speakers of Italian in Croatia. *Front. Commun.* 9:1371126. doi: 10.3389/fcomm.2024.1371126

The relationship between self-assessment of language proficiency and measures of lexical diversity and syntactic complexity: evidence from bilingual speakers of Italian in Croatia

Gordana Hržica[1], Sara Košutar[2]* and Nada Poropat Jeletić[3]

[1]Department of Speech and Language Pathology, University of Zagreb, Zagreb, Croatia, [2]Department of Language and Culture, UiT The Arctic University of Norway, Tromsø, Norway, [3]Department of Italian studies, Juraj Dobrila University of Pula, Pula, Croatia

A wide range of tools have been used to assess the language proficiency of bilingual speakers. The validity and high reliability of lexical diversity and syntactic complexity measures as instruments for measuring language proficiency have been demonstrated in previous studies across different languages. However, the relationship between self-assessment and the two measures has not yet been investigated. The present study focused on the Italophone bilingual language speakers, an understudied minority diglossic community in Croatia, and investigated whether measures of lexical diversity and syntactic complexity reflect self-assessment of language proficiency in the standard Italian language and the Istrovenetian dialect overall and in four specific domains (reading, writing, speaking, and listening). In addition, we aimed to investigate whether there are possible differences in self-assessment between the standard Italian language and the Istrovenetian dialect and whether there are language variety-related differences (standard vs. dialect) in the relationship between self-assessment and measures of lexical diversity and syntactic complexity. The results showed an intricate interplay between self-assessment and the lexical diversity and syntactic complexity of bilingual speech. This suggests that these measures are interrelated and that heritage bilingual language speakers may be able to objectively assess their language proficiency.

KEYWORDS

bilingualism, language proficiency, self-assessment of language proficiency, lexical diversity, syntactic complexity

Introduction

Bilingual speakers encounter considerable linguistic diversity in their daily communication, and the dynamics of these language experiences affect their language and cognitive functioning. Therefore, when considering the main constructs of bilingualism, a multidimensional approach is required taking into account a range of factors that have been shown to predict language performance in this population, for example language history,

amount and quality of language exposure, age of acquisition, formal education (Hoff et al., 2012; Rothman et al., 2023) and parental input (Unsworth et al. 2019; Romano and Sorace, 2023). There is a growing consensus on the need for an efficient and comprehensive instrument to assess the multidimensional construct of bilinguals' language experience (see Macbeth et al., 2022).

Two main approaches to measuring the language proficiency of bilinguals can be discerned: the external approach and the internal approach (Li and Zhang, 2021; Oscarson, 1989). The first approach involves language assessment through language tests, specific tasks, or teachers' assessment to gather data on language proficiency of a speaker, having in mind a predefined goal of assessment. The second approach is based on different forms of self-assessment in which individuals provide information about what they can do with the language or how far they have progressed. In line with the above classification, the two forms of assessment are also referred to as objective and subjective (Treffers-Daller, 2015). Subjective measures include questionnaires and self-ratings in a single language domain or multiple domains, i.e., speaking, listening, reading, and writing. Objective forms of assessment include, for example, test of different types that measure specific components of language proficiency or specific skills in one or/and the other language, such as fluency, general lexical knowledge, preferences, reading speed, etc. The latter forms of assessment also include elicited written production and transcripts of spoken language samples in spontaneous or narrative contexts (e.g., Montrul, 2015; Talamas et al., 1999; Treffers-Daller and Korybski, 2015).

In the last few decades, language bilingualism has increasingly attracted the attention of researchers (see Polinsky, 2018; Montrul and Polinsky, 2021). Bilingual speakers are a diverse group that can vary greatly in terms of their language history, language experiences, formal education in the languages they use as well as their language proficiency. From the perspective of language assessment research, particularly challenging and less researched are bilingual speakers living in communities with diglossia. Speakers of these communities receive input from both the dialect and the standard variety, navigating between the two language varieties depending on the context.

A wide range of tools, both subjective and objective, have been used to assess the language proficiency of bilingual speakers, both of which have been criticized in one way or another, mainly on the grounds that it is difficult to capture the complexity of bilingualism. The present study aims to contribute to this line of research by examining the relationship between self-assessment and objective measures of language proficiency (i.e., measures of lexical diversity and syntactic complexity), as well as possible language variety-related differences (regional dialect vs. standard) in this relationship. The particular focus of the study is on the language assessment of bilingual speakers of Italian living in Istria, a statutory bilingual county in Croatia, who are exposed to both the regional dialect (i.e., Istrovenetian) and standard Italian in their everyday lives, a phenomenon called *diglossia*. Istrovenetian (ISO 639-3: VEC) is a variety of Italo-Romance that belongs to the diatopic eastern branch of the Venetial dialectal system, alongside with Triestine and Dalmatian-Venetian. Historically, it was introduced during the colonial expansion of the Republic of Venice, supplanting the other Romance idioms spoken in the area and it became a regional koine' or *lingua franca*. It represents the mother tongue (or one of the mother tongues) of the members of the autochthonous Italian National Community in Croatia, characterized by a strong ethnolinguistic vitality within the Italophone communicative repertoire in Istria. Bilingual communities facing diglossia may assess their language proficiency in dialect and standard language differently. Therefore, the study of such bilingual communities provides a unique opportunity to investigate linguistic diversity among bilingual speakers.

Subjective forms of assessment: self-assessment

Self-assessments, in which bilinguals report on their language history, use and proficiency, are one of the most commonly used language assessment measures in the field of bilingualism and second language acquisition (SLA) (Treffers-Daller, 2015). There are a variety of self-assessment measures used to assess the linguistic background of bilinguals, for example the Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007) and the Language and Social Background Questionnaire (LSBQ; Anderson et al., 2018; Luk and Bialystok, 2013), which have been shown to be valid and reliable measures of the language background of bilinguals (Luk and Bialystok, 2013; Marian et al., 2007). To assess language proficiency, these questionnaires ask bilinguals to rate their current level of fluency in speaking, understanding (also referred to as listening or listening comprehension; see Brantmeier et al., 2012; Ross, 1998), reading, and/or writing in each of their languages.

There is empirical evidence that self-assessment is an appropriate and predictive measure of language proficiency (Delgado et al., 1999), both in terms of overall language and domain-specific proficiency (Delgado et al., 1999; Flege et al., 1999), which has been confirmed in numerous studies on bilingual and SLA research on adults and children with typical and impaired language development (i.e., Bedore et al., 2010), highly proficient (second) language learners (Clyne, 1997; Williams and Hammarberg, 1998) and less proficient ones (De Angelis, 2007; Selinker and Baumgartner-Cohen, 1995).

However, although self-assessment is a widely used method, its reliability, validity, and usefulness have been largely disputed (e.g., Blanche and Merino, 1989; Polinsky, 2018; Ross, 2006). Given the complexity of language use and proficiency in a range of tasks, some researchers doubt that self-assessment is detailed enough to capture this, even when specific domains of proficiency are considered (e.g., Treffers-Daller, 2015). Additionally, it has been found that self-assessments can be influenced by individuals' own biases, as individuals tend to overestimate or underestimate their language proficiency (e.g., Blanche and Merino, 1989; MacIntyre et al., 1997).

Therefore, self-assessment alone might not be considered the most appropriate instrument for language assessment, and other and more comprehensive quality indices are needed, possibly combining both subjective and objective measures and allowing a more in-depth investigation of the relationship between these measures (see Kang and Kim, 2016; Macbeth et al., 2022; Treffers-Daller, 2015). This encouraging approach is further developed in the present study, which aims to investigate the relationship between self-assessment and objective forms of assessment in bilingual speakers.

Objective forms of assessment: measures of syntactic complexity and lexical diversity

Objective forms of assessment often include language proficiency tests or teachers' assessments. Another way to objectively measure the language proficiency of bilingual speakers is to analyze spontaneous language production and gain an insight into current naturalistic language use (see Macbeth et al., 2022; Treffers-Daller, 2015). Language sample analysis (LSA) offers an ecologically valid, naturalistic, and efficient form of language assessment that can be used together with standardized language tests to assess bilingual language proficiency (see Ebert, 2020). However, compared to language tests, LSA has greater ecological validity and allows for a more in-depth analysis of different aspects of language that can be assessed in a single assessment task.

Two main concepts have been used as indices of language ability when applying LSA, syntactic complexity and lexical diversity. Syntactic complexity refers to the range of syntactic structures and the degree of sophistication of those structures (Ortega, 2003). It is a multidimensional construct that can be measured at the level of overall complexity, complexity via subordination, and subclausal complexity (Norris and Ortega, 2009). The overall syntactic complexity is usually operationalized by global or generic metrics of language complexity such as the mean length of communication unit (C-unit) (MLCU; Heilmann et al., 2010; Loban, 1976), mean length of utterance measured in words (MLUw; Bishop and Donlan, 2005; Frizelle et al., 2018), or mean length of turn (MLT; Kelly et al., 2022; Peltonen, 2021).

The most common global measure of syntactic complexity is MLCU. C-unit consists of a main clause with its modifiers or a main clause with all subordinate clauses attached to it. A clause is any syntactic unit consisting of at least one predicate. The length of the C-unit increases if it consists of a dependent clause or if the syntax within a clause is more complex, e.g., if a clause is extended by adding attributes, appositions, etc. More fine-grained measures of syntactic complexity include clausal density (CD), calculated as the total number of main and subordinate clauses divided by the total number of C-or T-units (Gutierrez-Clellen and Hofsteter, 1994; Mäkinen et al., 2014), or mean length of clause (MLC) as a more specific measure of syntactic complexity at the subclausal level. Measures of syntactic complexity have been used to assess language abilities in first language acquisition (e.g., Košutar et al., 2022), language proficiency of L2 learners (e.g., Ortega, 2003; Peltonen, 2021), as well as the language abilities of bilingual children (e.g., Andreou and Tsimpli, 2020).

Mean length of turn (MLT; Caissie et al., 1998) measures the length of turn-taking in the conversation and the division of conversational load between different speakers. It is measured by the number of words or number of syllables per speaking turn (e.g., Nitta and Nakatsuhara, 2014) and provides information about the control of the conversational situation. Indirectly, it can point to different aspects of language proficiency of a speaker. In child language development research, MLT has been used as a measure of language complexity (e.g., Creaghe, 2019; Ninio and Snow, 1999). In the L2 studies MLT has been used to measure fluency of L2, both in the broader sense (fluency as oral proficiency) or narrower, the flow and smoothness of speech (van Os et al., 2020). A higher MLT suggests longer and more elaborate contributions from speakers, potentially indicating more complex or detailed storytelling, explanations, or discussions. On the other hand, a lower mean length of turn may point to shorter and more concise utterances, possibly signaling a more rapid flow of conversation. In general, MLT provides information about the length of utterances and larger sections of connected speech, which can reflect the complexity of syntax (Fagan and Iglesias, 2000; Van der Veen et al., 2021).

Lexical diversity refers to how diverse the vocabulary produced is. The more diverse the vocabulary, the greater the lexical diversity. Traditional measures of lexical diversity include number of different words (NDW; Miller, 1981) and type-token ratio (TTR; Templin, 1957). TTR is calculated as the total number of unique words (types) divided by the total number of words (tokens). Although these measures have been widely used, both have been shown to be influenced by the language sample's length (Malvern et al., 2004). To overcome these limitations, researchers proposed alternative measures that consider text length, such as measure D (Malvern and Richards, 1997) and moving-average type-token ratio (MATTR, Covington and McFall, 2010). Measure D is based on mathematical modelling of the decreasing TTR curve with the increasing length of the language sample. MATTR calculates the TTR of text windows with a fixed size by moving the window through the text. At the end of the text, all TTRs are averaged to determine the final score. Both D and MATTR provide valid assessments of lexical diversity even when language samples differ in length (Fergadiotis et al., 2015 for MATTR; Jarvis, 2002; McCarthy and Jarvis, 2010 for D). Measures of lexical diversity have been used to assess language abilities of bilingual children (e.g., Mitrofanova et al., 2018) and L2 learners (Treffers-Daller et al., 2018). Previous studies on Croatian speakers have also found that measure D can predict receptive vocabulary scores in bilingual children (e.g., Hržica and Roch, 2021).

The relationship between self-assessment and objective forms of assessment

Since self-assessment is a relatively simple and learner-centered, it has long been of great interest to investigate the validity of self-assessment measures by testing the relationship between self-assessment and objective forms of assessment to find out whether the content and construct measured by the two forms of assessment are identical (e.g., Bachman, 1990; DeVellis, 2003; Ma and Winke, 2019). In general, positive correlations were found, but the strength of the correlation varied. In their meta-analysis, Li and Zhang (2021) found an average aggregate correlation of.466 between self-assessment and objectively measured language performance (i.e., language tests and teacher assessment). However, they also discovered that there are factors (e.g., criteria type of self-assessment, training, total numbers of items in the self-assessment instrument) that weaken this relationship or contribute to the strength of the correlation coefficient.

Self-assessment can be measured as an overall score or more specifically for different language domains, i.e., listening, speaking, reading, and writing. As Zell and Krizan (2014) have shown, correlations between self-assessment and objective measures are stronger if the self-assessment is calculated specifically for a specific domain and not as an overall proficiency. There is evidence that self-assessment of receptive language skills (i.e., listening and reading) correlates more strongly with language tests than self-assessment of productive skills (i.e., speaking and writing) (Krausert, 1991;

Ross, 1998). In their meta-analysis, Li and Zhang (2021) found the highest correlation between self-assessment of listening and the language tests, followed by reading and speaking and a significantly lower coefficient for self-assessment of writing, although all these correlations were significant and moderate.

The studies to date differ greatly in the methodology used, especially in the self-assessment procedure. It seems that greater specificity and explicitness of self-assessment items strengthens the relationship between self-assessment measures and objective forms of assessment (Andrade and Valtcheva, 2009; Panadero and Romero, 2014) and the same is true when speakers are trained in advance (Birjandi and Bolghari, 2015; Ross, 2006). Both specificity and training were significant factors in this regard (Li and Zhang, 2021).

More consistent results have been obtained when the relationship between language tests and self-assessment measures has been examined [moderate to high correlations in Bachman and Palmer (1989) and Ikeguchi (1996)], while the relationship between teacher assessment and self-assessment is not so clear. Birjandi and Bolghari (2015) reported high correlations, while Langan et al. (2008) found weak or non-significant correlations. In their meta-analysis, Li and Zhang (2021) found that both language tests and teacher assessments were moderately correlated with self-assessment scores. The correlation was slightly lower for teacher assessments, but the discrepancy was not statistically significant.

Previous research has shown that there is a relationship between self-assessment and objective measures of language performance (e.g., language tests and teachers' assessments) and that this relationship may depend among others on the type of objective measure used. As far as we know, there is no study that has investigated the relationship between self-assessment and language sample measures, namely objective forms of language assessment, such as measures lexical diversity and syntactic complexity.

Challenges in the language assessment of bilingual speakers

Due to the ability to use their languages in different contexts and with different interlocutors, bilinguals have a variety of language experiences in their everyday lives. In assessing their language proficiency, researchers have used a range of tasks, both subjective and objective (e.g., Hayakawa et al., 2022; Macbeth et al., 2022; Montrul, 2011; Polinsky, 2018); nevertheless, self-assessments have often been criticized, especially in studies on heritage bilingual speakers. Heritage speakers are individuals who acquire the minority language (i.e., their heritage language) at home and the majority language (i.e., the community language) through immersion during childhood, often after starting formal education (Montrul, 2015; Rothman, 2009). When measuring the language proficiency of bilingual heritage speakers, it can be a challenge to determine the baseline language to be assessed. The baseline language is not necessarily the standard variety of the native-speaking population or the variety that is taught in formal education. In most cases, the home language of the heritage speaker is a regional dialect, and the exposure to other dialects or a formal standard variety is not that common (see Polinsky, 2014). In communities where dialects are the primary means of communication and speakers have limited exposure to the standard variety, language assessment on these dialects is crucial for understanding the dynamics of the linguistic reality of the community. Dialects in heritage language communities are integral parts of linguistic identity and reflect the unique history, culture, and social dynamics of a particular group (see Brehmer, 2021).

Polinsky (2018) challenges the use of self-assessments by heritage speakers by pointing out that heritage speakers' perceptions may differ inside and outside of their home speech community. Some heritage speakers, who are criticized because of the way they speak, may compare themselves negatively with baseline speakers, but those same speakers may show a different attitude towards their heritage language in the context where another language is dominant. Moreover, some researchers suggest that self-assessment by bilingual speakers may be inversely correlated with their language proficiency (e.g., Beaudrie and Ducar, 2005; Davidson and Lekic, 2013; Thompson, 2015; Titus, 2012). This is likely due to the fact that the higher the speaker's language proficiency, the more aware they are of their linguistic limitations. Indeed, Polinsky (2018) has reported a negative correlation between self-assessment and objective assessment, namely fluency measured as speakers' speech rate, in Russian heritage speakers. The more fluent the speakers are, the greater their meta-linguistic awareness and the lower their self-esteem.

Contrary findings were reported in other studies on bilingual heritage speakers. For example, Macbeth et al. (2022) have found self-reported heritage language use to be moderately and positively related to self-reported overall proficiency of these speakers (i.e., self-reported proficiency ratings for speaking, reading, writing, and listening). The more a heritage language speaker used their heritage language, the more proficient they report themselves to be in their heritage language. Moreover, this study has confirmed a relationship between self-assessment and objective assessment in bilingual heritage speakers. The results from the laboratory test of language proficiency (i.e., verbal fluency test) were positively related to and significantly predicted the results from self-reported overall proficiency, suggesting that a laboratory-based proficiency is consistent with how heritage speakers self-report their overall proficiency. On the other hand, spontaneous language (speech) use did not significantly predict self-reported overall proficiency, but rather self-reported heritage language use. Ultimately, none of the three forms of assessment correlated strongly with each other and each provided unique information about the heritage bilingual language experience. Although Macbeth et al. (2022) have confirmed the relationship between self-report by heritage speakers and objective language assessment, the authors have only investigated the overall self-reported proficiency of these speakers and did not provide any insights into the relationship between self-reported proficiency in a particular domain and different forms of objective assessment (i.e., language proficiency tests, spontaneous language production).

Kang and Kim (2016) have also found close relationship between self-assessment (i.e., writing skills, oral fluency, vocabulary, pronunciation, and grammar) and objective assessment (i.e., the amount and quality of output in speaking and writing) of language skills of Korean heritage speakers living in America. However, the results for the two forms of assessment were not identical, leading the authors to suggest that self-assessment would be better used as a complementary assessment tool in bilingual heritage speakers. In addition, bicultural identity was a strong predictor of self-assessment of speaking and writing in heritage speakers. Similar results were reported in another study by Kang and Kim (2012). These findings

suggest that the degree of cultural identity of speakers can contribute to their self-assessment of their language proficiency (see also Macbeth et al., 2022). Speakers with a stronger cultural identity tend to self-rate their heritage language proficiency higher.

Furthermore, language assessment might be particularly challenging in speakers of bilingual communities where, sometimes, two varieties of a language coexist within a community (a phenomenon called *diglossia*), thus further delineating the spectrum of bilingual language experience (Milani Kruljac, 2001, 2003). Such communities are driven by community initiatives, educational programs (i.e., from primary schools to universities), government policy and media, which provides a more conducive environment for the promotion of heritage language and cultural preservation (i.e., Lanthaler, 2001). The standard variety is generally used in education, in the media and in certain situations (e.g., in official documents), while the dialect is common in everyday communication (e.g., within families and local communities) (i.e., Blagoni et al., 2016). This duality adds further layers to the linguistic dynamics within the community, and speakers navigate this multi-layered linguistic reality by balancing between the use of standard variety and dialect depending on the context (see Iannaccaro et al., 2003). An example of such a bilingual community are the speakers of Italian in Istria, a statutory bilingual county in Croatia, who are exposed to both the regional dialect (i.e., Istrovenetian) and standard Italian in their everyday lives.

Therefore, when assessing the language proficiency of bilingual speakers living in such communities, one should consider both language varieties used by the speaker. Such an approach can contribute to a more comprehensive understanding of the diverse experiences that heritage bilinguals encounter. Moreover, self-assessment in specific domains (i.e., speaking, listening, reading, and writing) may vary depending on which language variety the speaker is more familiar with and may therefore relate differently to different forms of objective assessment (e.g., tests of language proficiency, languages samples). To the best of our knowledge, no study has yet looked at the possible differences in self-assessment between standard variety and dialect or provided insights into possible language variety-related differences in the relationship between self-assessment and objective language assessment of bilingual speakers.

Research questions

A wide range of subjective and objective tools have been used to assess the language proficiency of bilingual speakers, all of which have been criticized in one way or another (see Blanche and Merino, 1989; Ross, 2006). Despite criticism, self-assessment has been shown to be a reliable and valid tool for assessing the language proficiency of bilingual speakers (Delgado et al., 1999), and it has also been recommended to be used in combination with objective measures (Kang and Kim, 2016). However, previous studies have shown contradictory results regarding the relationship between self-assessment and objective forms of assessment (*cf.* Macbeth et al., 2022; Polinsky, 2018). Moreover, they did not provide information on the relationship between objective assessment and self-assessment in specific domains (i.e., listening, reading, speaking, and writing). Possible differences related to the linguistic diversity that bilingual speakers are exposed to (i.e., standard vs. dialect) have also not been investigated. As a result, there is a lack of knowledge about language assessment in different language varieties, and about the possible language variety-related differences in the relationship between the two forms of assessment.

The present study aims to contribute to this line of research by investigating the relationship between self-assessment and objective form of assessment such as language sample analysis (LSA), i.e., measures of syntactic complexity and lexical diversity in spontaneous speech production. The particular focus of the study is on the language assessment of Italian bilingual speakers living in Istria, a statutory bilingual county in Croatia, who are exposed to both the regional dialect (i.e., Istrovenetian) and standard Italian in their everyday lives. We were also interested in investigating the potential language variety-related differences (standard vs. dialect) in the relationship between the two forms of language assessment to gain a more comprehensive understanding of the language proficiency of bilingual speakers.

Based on the results of previous studies, we addressed the following research question: Is there a relationship between self-assessment (overall and in specific domains: listening, speaking, reading, and writing) and measures of lexical diversity and syntactic complexity in the standard Italian and the Istrovenetian dialect? If so, how can it be described?

Materials and methods

Participants

The participants in this study are bilingual speakers of Croatian and Italian who live in the bilingual region of Istria County in Croatia. Istria County is located in the north-western part of Croatia and is a border region that has been in constant contact with Slavophone and Italophone cultures and languages for several centuries. Apart from the Croatian majority community, the most numerous minority community and the only recognized national minority in Istria County is the autochthonous Italian National Community (e.g., Blagoni et al., 2016; Milani Kruljac, 2001, 2003; Giuricin and Giuricin, 2008). Istria has a history of enduring and stable societal bilingualism, with Croatian-Italian bilingualism being officially recognized (both languages are official in the County). Furthermore, a significant part of the Istrian population is bilingual, with the Italophone-speaking community being (almost) completely bilingual today.

The selected group comprises 82 participants who are Italophone bilingual speakers of Italian (Istrovenetian) and live in Istria ($N = 82$; $M = 28$, $F = 54$; age range = 20–50+). Participants reported speaking both Italian Istrovenetian and Croatian, with Italian Istrovenetian being the mother tongue for the majority (68%). All participants began acquiring both languages at an early age, with the second language introduced by the latest at age 3. The majority of participants were female (66%), and most were in their twenties (see Table 1).

TABLE 1 Information about participants.

N	82	
Mother tongue	Italophone speakers	Croatophone speakers
	68%	32%
Age	Mean	Range, SD
	39	16–80, 15.45

Materials

The participants' data were extracted from the larger data pool of the *Corpus of Spoken Istrovenetian/Fiuman and Croatian* (C-ORAL-IC) (Poropat Jeletić et al., 2024), which contains 41 transcripts of conversations, corresponding source audio files (37 media files) and an accompanying participant spreadsheet (91 participants) with demographic and sociolinguistic data on each speaker. The data was collected between 2018 and 2021. Language sampling was conducted by researchers from bilingual communities who had access to groups of bilingual speakers. Sampling took place in various daily, informal, and interactive situations, especially in spontaneous speech situations among family members, friends, colleagues, or acquaintances, such as informal gatherings, socializing or family dinners. As speakers were recorded in informal situations, they spoke Italian dialect, namely Istrovenetian. All speakers gave a written consent form in which they agreed to be recorded without their explicit knowledge within 1 month of signing the consent form. Investigators were trained to minimize their involvement in the recorded sessions to avoid the Observer's paradox (1972) as much as possible. The corpus was transcribed using Talk Bank's uniform transcription standard (CHAT) as the coding system (Mac Whinney, 2000). The participants spoke Italian most of the time. Utterances and words in Croatian were marked so that they could be excluded from the analyses. For this study, we selected all participants with more than 80 utterances.

The data on self-assessment of language proficiency was collected using a bilingual questionnaire. The questionnaire is available on the Open Science Framework (OSF): https://osf.io/mkg72/. Each participant was asked to rate their language proficiency in the standard Italian and in the Istrovenetian dialect in four domains: listening, speaking, reading, and writing. A five-point Likert scale was used for this purpose. Participants answered these questions as part of a broader online survey, which also included questions about their age, language status, education, etc. The survey was distributed via the online platform SurveyMonkey.

We used the *Computerized Language Analysis program* (CLAN; MacWhinney, 2000) to calculate measures of syntactic complexity and lexical diversity for each participant. Measures of syntactic complexity included mean length of communication unit (Heilmann et al., 2010) and mean length of turn (MLT; Caissie et al., 1998). MLCU was calculated in words using CLAN's MLU program, which divides the number of words (tokens) spoken by a participant by the number of lines uttered by that participant in the transcripts. To segment the spontaneous speech into utterances, communication units (C-units) were used. Each line of the transcript represents one C-unit, which was segmented according to syntactic criteria. C-units are defined as independent clauses with their modifiers. Main clauses that are independent can be segmented into one C-unit. However, subordinate clauses that are dependent on the main clause cannot stand alone. Therefore, the C-unit consists of either a main clause or a main clause with its subordinate clauses. MLT was calculated as the average length of a turn in words using the CLAN's program MLT. This program divides the number of words spoken by a participant by the number of rounds of conversation that participant has completed. For both measures of syntactic complexity, utterances in Croatian were excluded from the analysis, but the Croatian words contained in the utterance were counted.

For lexical diversity, D measure was chosen, as it has been shown to be relatively independent of the size of the language sample (Malvern et al., 2004). The measure was calculated using the CLAN's program VOCD, excluding utterances and words in Croatian, as well as repetitions, self-corrections, and other disfluencies (e.g., filled pauses).

Data analysis

Since the data were not normally distributed, we used the non-parametric Spearman's correlation to assess the correlation between the self-assessment measures (listening, speaking, reading, writing) in standard Italian and Istrovenetian dialect and the language sample measures (MLCU and D). We also calculated the average self-assessment score and determined Spearman's correlation between this score and the language sample measures. The analyses were performed using SPSS version 20.0. (IBM IBM Corp, 2011).

Results

Descriptive statistics

Measures of lexical diversity and syntactic complexity in Istrovenetian

Scores for lexical diversity and syntactic complexity are reported in the Table 2. Speakers talked in Italian dialect (Istrovenetian).

In terms of lexical diversity, the average score for measure D is 92.4, with individual scores ranging from a minimum of 35.49. However, a large number of participants have a lexical diversity score of over 60. Similar variability can be observed when measuring the MLCU. The average number of words per communication unit is 5, with the lowest value being 1.69. Despite this range, the majority of participants consistently produce C-units with an average length of more than 3.5 words. The mean length of turn (MLT) also shows great variability, ranging from only 3 words to 28 words. Most participants produced turns with an average length of 9 words or more.

Self-assessment scores in standard Italian and Istrovenetian

Scores for self-assessment in Table 3 for Italian and in Table 4 for Istrovenetian.

The participants rated their language proficiency in standard Italian with consistently high scores. The average scores for listening, speaking, reading, and writing were above 4.6, with the highest average for listening and the lowest for writing. Whilst there was some variability between participants, illustrated by the minimum scores

TABLE 2 Descriptive results for measures of lexical diversity and syntactic complexity.

	Min	Max	*M*	*SD*
D (lexical diversity)	35.49	118.93	92.4	26.89
MLCU (syntactic complexity)	1.69	8.49	5	1.41
MLT (in words)	2.82	27.65	9.08	4.75

D, measure D; MLCU, mean length of communication unit; MLT, mean length of turn (in words).

TABLE 3 Descriptive results for self-assessment of standard Italian language.

	Min	Max	*M*	*SD*
Listening	3	5	4.82	0.45
Speaking	2	5	4.68	0.63
Reading	1	5	4.73	0.70
Writing	2	5	4.67	0.65

TABLE 4 Descriptive results for self-assessment of Istrovenetian.

	Min	Max	*M*	*SD*
Listening	1	5	4.73	0.45
Speaking	1	5	4.60	0.63
Reading	1	5	4.46	0.70
Writing	1	5	4.38	0.65

within each category, the lowest minimum scores were observed for reading and the highest for listening.

The assessment of language skills in the Istrovenetian dialect also resulted in high average scores, albeit slightly lower than for standard Italian. Nevertheless, all scores were above 4.3, with the highest average for listening and the lowest for writing. It is worth noting that the lowest score for each of the four language skills was 1.

Relationship between self-assessment and measures of lexical diversity and syntactic complexity in standard Italian and Istrovenetian

Correlations between self-assessment scores in standard Italian and Istrovenetian are presented in Table 5. First, we explored how self-assessment scores in different domains within one language varieties correlate. The overall self-assessment scores for standard Italian and the dialect showed a moderate correlation [$r_s(82) = 0.612$, $p < 0.001$]. Self-assessment scores for different domains of standard Italian all show significant high correlations (>0.7). Self-assessment scores for different domains of Istrovenetian dialect also show significant high correlations (>0.7), except for writing and listening, which show a moderate correlation [$r_s(82) = 0.682$, $p < 0.001$]. Next, we explored correlations between two varieties. The overall self-assessment scores for standard Italian and the Istrovenetian dialect showed a moderate correlation [$r_s(82) = 0.612$, $p < 0.001$], which is also the case for the individual domains. Moderate positive correlations were found between listening in standard Italian and in the dialect [$r_s(82) = 0.545$, $p < 0.001$], as well as between self-assessments of speaking [$r_s(82) = 0.511$, $p < 0.001$], reading [$r_s(82) = 0.539$, $p < 0.001$], and writing [$r_s(82) = 0.550$, $p < 0.001$]. Based on the observed positive, but lower correlations we conclude that it is appropriate to treat self-assessment in standard Italian and self-assessment in dialect as separate measures, as is also the case for the individual domains.

Correlations between self-assessment scores and measures of lexical diversity and syntactic complexity are presented in Tables 6, 7, separately for self-assessments in standard Italian and the Istrovenetian dialect. Self-assessment scores in the different varieties were compared with the unique values of indices for lexical diversity and syntactic complexity obtained from the language sample in the Istrovenetian dialect. For the standard Italian self-assessment scores, a low positive correlation was found between the overall self-assessment score and MLT [$r_s(82) = 0.249$, $p < 0.05$], while there was no significant correlation between self-assessment and either the measure D or MLCU. However, positive correlations were observed between specific domains of self-assessment in standard Italian and the measure D. Significant low correlations were found between the self-assessment of listening and D [$r_s(75) = 0.228$, $p < 0.05$] and between the self-assessment of reading and D [$r_s(75) = 0.242$, $p < 0.05$]. The self-assessment of writing did not correlate with D. There were no observed correlations between specific domains of self-assessment in standard Italian and syntactic complexity measured by MLCU. Positive correlations were observed between specific domains of self-assessment in standard Italian and syntactic complexity measured by MLT. MLT correlated positively with listening [$r_s(82) = 0.274$, $p < 0.05$] and reading [$r_s(82) = 0.303$, $p < 0.01$], but not with speaking [$r_s(82) = 0.201$, $p = 0.07$] and writing [$r_s(82) = 0.160$, $p = 0.153$].

For the Istrovenetian dialect, the overall self-assessment score showed a significant positive low correlation with MLT [$r_s(82) = 0.256$, $p < 0.05$] and the measure D [$r_s(82) = 0.242$, $p < 0.05$], but not with MLCU. Correlations were found for specific domains of self-assessment. There were significant positive low correlations between self-assessment of listening and D [$r_s(75) = 0.234$, $p < 0.05$] and between the self-assessment of writing and D [$r_s(75) = 0.239$, $p < 0.05$]. A marginally significant low positive correlation was observed between the self-assessment of speaking and D [$r_s(75) = 0.213$, $p = 0.066$], while no significant correlation was found between the self-assessment of reading and D. There were no observed correlations between specific domains of self-assessment in standard Italian and syntactic complexity measured by MLCU. In particular, there were no significant correlations between the self-assessment of listening and MLCU [$r_s(82) = 0.158$, $p = 0.155$], the self-assessment of speaking and MLCU [$r_s(82) = 0.099$, $p = 0.378$], the self-assessment of reading and MLCU [$r_s(82) = 0.196$, $p = 0.077$], and the self-assessment of writing and MLCU [$r_s(82) = 0.057$, $p = 0.612$]. Self-assessment scores in some domains for Istrovenetian dialect showed correlation with measures of syntactic complexity. A significant positive low correlation was found between the self-assessment of listening and MLCU [$r_s(82) = 0.217$, $p = 0.05$], but not for any other domain. Positive correlations were observed between specific domains of self-assessment in standard Italian and syntactic complexity measured by MLT. The overall scores for self-assessment showed a significant low positive correlation with MLT [$r_s(82) = 0.249$, $p < 0.05$]. Regarding the self-assessment in specific domains, MLT correlated positively with listening [$r_s(82) = 0.274$, $p < 0.05$] and reading [$r_s(82) < 0.303$, $p < 0.01$], but not with speaking [$r_s(82) = 0.201$, $p < 0.07$] and writing [$r_s(82) = 0.160$, $p < 0.153$] MLT showed a significant positive correlation with all domains of self-assessment: listening [$r_s(82) = 0.322$, $p < 0.01$], speaking [$r_s(82) = 0.257$, $p < 0.05$], reading [$r_s(82) < 0.296$, $p < 0.01$], and writing [$r_s(82) = 0.249$, $p < 0.05$].

Discussion

The present study aimed to investigate the relationship between self-assessment of language proficiency and objective measures of lexical diversity and syntactic complexity in bilingual speakers of Italian living in Croatia, taking into account assessment in two

TABLE 5 Correlations between self-assessment scores in different language varieties (standard Italian vs. Istrovenetian dialect).

	1	2	3	4	5	6	7	8	9	10
Listening (SI) – 1	–									
Speaking (SI) – 2	0.794**	–								
Reading (SI) – 3	0.959**	0.752**	–							
Writing (SI) – 4	0.753**	0.775**	0.799**	–						
Overall (SI) – 5	0.832**	0.875**	0.866**	0.909*	–					
Listening (ID) – 6	0.545**	0.472**	0.521**	0.407**	0.509**	–				
Speaking (ID) – 7	0.520**	0.511**	0.492**	0.386*	0.480**	0.844**	–			
Reading (ID) – 8	0.561**	0.598**	0.539**	0.507**	0.599**	0.764**	0.806**	–		
Writing (ID) – 9	0.475**	0.535**	0.502**	0.550**	0.592**	0.682**	0.760**	0.878**	–	
Overall (ID) – 10	0.513**	0.571**	0.528**	0.533**	0.612**	0.729**	0.811**	0.912**	0.991**	–

SI, standard Italian; ID, Istrovenetian dialect. **, correlation is significant at the 0.01 level (two-tailed). *, correlation is significant at the 0.05 level (two-tailed).

TABLE 6 Correlations between measures of lexical diversity, syntactic complexity, and self-assessment scores in Italian standard language.

	1	2	3	4	5	6	7	8
D – 1	–							
MLCU – 2	0.129	–						
MLT – 3	0.273*	0.803**	–					
Listening – 4	0.228*	0.158	0.274*	–				
Speaking – 5	0.076	0.099	0.201	0.794**	–			
Reading – 6	0.242*	0.242*	0.303*	0.752**	0.959**	–		
Writing – 7	0.143	0.143	0.160	0.753**	0.775**	0.799**	–	
Overall – 8	0.125	0.153	0.249*	0.832**	0.875**	0.866**	0.909**	–

D, measure D; MLCU, mean length of communication unit; MLT, mean length of turn (in words). **, correlation is significant at the 0.01 level (two-tailed). *, correlation is significant at the 0.05 level (two-tailed).

language varieties, the standard Italian language and the Istrovenetian dialect.

The results of the descriptive statistics for self-assessment in all four domains in the two language varieties are consistently at or above 4, which indicates that the participants consider themselves to be proficient speakers of both the standard Italian and the Istrovenetian dialect. However, while the scores for standard Italian are consistent in all four domains and are above 4.5, there is a difference in the scores for the self-assessments of the Istrovenetian dialect. Self-assessment scores are higher for listening and speaking, when compared to reading and writing. These results are consistent with the self-assessment results of heritage speakers of Italian in Romano and Guijarro-Fuentes (2023) and findings for heritage speakers of Spanish in the USA, who are known to have advantages over L2 speakers in tasks involving listening and speaking but are at a disadvantage in tasks involving reading and writing (Bowles, 2011).

In the case of the present study, the observed dissociation between speaking and writing reflects that the dialect is primarily used in spoken communication as a modality which is inherently less complex than written language (e.g., Biber, 1988; Roland et al. 2007). The complexity of written language stems from greater syntactic complexity usually associated with advanced writing, but also from the three components of vocabulary, namely lexical density, lexical sophistication, and lexical diversity (Durrant et al., 2021; Lu, 2012; Read, 2000). Crucially, the divergence in scores across the four domains for the Istrovenetian dialect, but not for the standard Italian, mirrors the functional differences of the two varieties. The standard Italian is a language variety used in formal contexts, such as educational contexts, while the dialect is predominantly used in everyday informal communication. This duality of language usage emphasizes not only the linguistic diversity but the sociolinguistic communication practices of this bilingual community. We examined the relationship between the overall scores for self-assessment and objective assessment, i.e., measures of lexical diversity and syntactic complexity. The overall score in the standard Italian positively correlated only with the measure MLT for syntactic complexity, while the overall score in the Istrovenetian dialect correlated positively with both the measure D for lexical diversity and the MLT for syntactic complexity. Previous research examining the relationship between self-assessment and objective forms of assessment obtained different results, ranging from weak correlations or non-significant results (e.g., Langan et al., 2008) to significant positive correlations (e.g., Bachman and Palmer, 1989; Li and Zhang, 2021). Moreover, it has been established that the nature of objective assessment is likely to influence the magnitude of correlation, for example stronger correlations were obtained for language tests than for teachers' assessments (Li and Zhang, 2021). Our results contribute to this existing body of evidence, but it is worth noting that the correlations we obtained are relatively low (*cf.* Li and Zhang, 2021). This could be the result of the methodology used. Notably, we employed a standard form of

TABLE 7 Correlations between measures of lexical diversity, syntactic complexity, and self-assessment scores in Istrovenetian dialect.

	1	2	3	4	5	6	7	8
D – 1	–							
MLCU – 2	0.129	–						
MLT – 3	0.273*	0.803**	–					
Listening – 4	0.234*	0.217*	0.322**	–				
Speaking – 5	0.213	0.181	0.257*	0.884**	–			
Reading – 6	0.150	0.192	0.296**	0.764**	0.806*	–		
Writing – 7	0.239*	0.144	0.249*	0.682**	0.760**	0.878**	–	
Overall – 8	0.242*	0.152	0.256*	0.729**	0.811**	0.912**	0.991**	–

D, measure D; MLCU, mean length of communication unit; MLT, mean length of turn (in words). **, correlation is significant at the 0.01 level (2-tailed). *, correlation is significant at the 0.05 level (two-tailed).

self-assessment, where speakers self-rated their language proficiency without providing them with more explicit criteria. Participants were asked to provide a score without explanations or examples that might help them to determine the appropriate score. This contrasts with alternative self-assessments that provides guidelines for scoring (e.g., Brown and Harris, 2013). Additionally, we utilized a Likert scale as opposed to a computer-assisted adaptive instrument for self-assessment. It is important to note that the explicit task criteria and the use of computer-assisted instruments have emerged as factors contributing to the strength of correlation in previous studies (Li and Zhang, 2021).

Next, we examined the relationship between the overall scores of self-assessments in specific domains (listening, reading, speaking, and writing) and objective measures of lexical diversity (D) and syntactic complexity (MLT, MLCU). The self-assessment of listening correlated with measure D both in standard Italian and Istrovenetian dialect, while listening in Istrovenetian dialect correlated with measure MLT. On the other hand, self-assessment of reading in standard Italian correlated moderately with D and MLT, while reading in Istrovenetian dialect correlated only with MLT. Next, only self-assessment of speaking in Istrovenetian dialect showed a significant correlation with MLT. Self-assessment of writing correlated with both D and MLT, but the relationship was significant only in Istrovenetian dialect.

Previous studies indicate that self-assessment of receptive language skills (i.e., listening and reading) shows a closer relationship with objective assessment than self-assessment of productive skills (i.e., speaking and writing) (e.g., Krausert, 2013; Ross, 1998). The present study confirms this relationship, revealing the robust (albeit low) correlations between listening and reading as receptive language skills and objective measures. A meta-analysis conducted by Li and Zhang (2021) underscores a notable correlation between self-assessment of listening and the language tests, aligning with our findings. Consequently, it appears that the relationship between objective measures and specific domains of self-assessment is more pronounced in the listening domain.

The obtained results reveal the intricate interplay between self-assessments and objective assessment, but also offers insights into the patterns of use of different language varieties in bilingual speakers.

We observed a significant correlation between overall self-assessment of language proficiency in standard Italian and the Istrovenetian dialect. This correlation is not very high, which indicates that speakers recognize differences in their language proficiency between the two language varieties. Such awareness is expected in a bilingual community with diglossia. Speakers are likely to have similar language proficiency in standard Italian and the Istrovenetian dialect to a certain extent. However, differences arise due to the educational background. The presence of both Croatian and Italian schools in the region as well as the different educational backgrounds of the participants (some participants received formal education in Italian, others in Croatian and a few in both languages) contribute to these differences. In addition, some participants attended universities in Italy or Croatia, which could have a further influence on the perception of their language proficiency. Overall, the observed differences in self-assessment between standard variety and dialect can be attributed to these differences in education.

Secondly, a greater number of correlations between the overall self-assessment and the measures of syntactic complexity and lexical diversity were found for the Istrovenetian dialect than for Standard Italian. Differences were also found in correlations between domain-specific self-assessments and measures of lexical diversity and syntactic complexity. We observed more significant correlations between domain-specific self-assessments and measures of lexical diversity and syntactic complexity in Istrovenetian dialect than in standard Italian. Higher number of correlations between self-assessment and objective measures in the dialect rather than in the standard Italian can be expected, as the participants were recorded in their everyday communication in which they primarily use the dialect. It is noteworthy that self-assessment results correlate with MLT in all domains, suggesting that better language proficiency contributes to longer conversational turns, which is consistent with previous studies (e.g., Creaghe, 2019; Ninio and Snow, 1999). These results emphasize the multi-layered nature of this relationship, which is particularly evident in the correlations across all domains of self-assessment within the dialect.

Higher number of correlations between self-assessment and objective measures in the dialect rather than in the standard variety can be expected, as the participants were recorded in their everyday communication in which they primarily use the dialect. It is noteworthy that self-assessment results correlate with MLT in all domains, suggesting that better language proficiency contributes to longer conversational turns, which is consistent with previous studies (e.g., Creaghe, 2019; Ninio and Snow, 1999). These results emphasize the multi-layered nature of this relationship, which is particularly evident in the correlations across all domains of self-assessment within the dialect.

In the present study, we investigated Italophone bilingual speakers in Croatia, who are still an under-researched and under-represented bilingual community. The results point to the importance of self-assessment on different language varieties to which bilingual speakers living in communities with diglossia are exposed. Different language varieties can have a different status and function. Self-assessment in specific domains (i.e., speaking, listening, reading, and writing) may vary depending on which language variety the speaker is more familiar with, and may therefore relate differently to different forms of objective assessment. For example, self-assessment in standard variety may be strongly related to language proficiency tests, while self-assessment in dialect may be more strongly related to the measures of language sample analysis. To the best of our knowledge, no study has yet examined the possible differences in bilingual language assessment between standard variety and dialect or language variety-related differences in the relationship between self-assessment and objective language assessment of bilingual speakers.

The obtained results are informative not only from the perspective of bilingual speakers of Istrovenetian in Croatia but can also be extended to other similar bilingual communities. For example, there are several other bilingual language communities that share the feature of *diglossia*, having in common similarities with the Istrian Italophone diglossic community: the German-speaking community in the north Italian autonomous province of South Tyrol (Lanthaler, 2001), where standard German if spoken in formal settings and the Tyrolean varieties in informal daily communication; the Catalan-speaking community in Catalonia (Miller and Miller, 1996), the Occitan-speaking community in South France (Blanchet and Schiffman, 2004), the Romansh-speaking language group in Graubünden in Switzerland (Regula, 2008), the Franco-Provençal speakers in Aosta Valley in Italy (Iannaccaro et al., 2003), among others. Although each of these language communities has a unique historical and cultural context, they face rich and diverse linguistic experiences, identity negotiations and efforts to preserve language and culture. Such communities experience a clear division between formal and informal settings that imply the use of diverse linguistic varieties (standard and dialect), while communities without diglossia exhibit a more flexible language use across different communicative contexts. The presence of diglossia thus characterizes the linguistic landscape and the dynamics of everyday communication.

There are some limitations of the present study, one of them being related to the fact that the participants largely identify themselves as native speakers of Italian. As members of a national minority, they may resemble to heritage language speakers, who often perceive their heritage language as their native language due to emotional ties, family, and cultural identity, despite their varying language proficiency (Rothman, 2009; He, 2010; Montrul, 2015). This could influence self-assessment and lead to an overestimation of language proficiency, namely heritage speakers might be more proficient in certain domains (e.g., family settings) but not others. Their language use varies based on social networks and environments (Montrul, 2015), and recent context-specific experiences may not reflect overall ability. All these factors may limit the reliability of self-assessment in heritage speakers and explain the low correlations in this research. However, the participants in our study live in an officially bilingual community with educational and cultural networks in both languages. Therefore, they do not fit the typical definition of heritage language speakers, as the language they consider to be their native language has a similar status to other languages in society (*cf.* Rothman, 2009). They might be less influenced by cultural and emotional factors and better able to objectively assess their language proficiency. Nevertheless, further research is needed on how language identity affects language assessment in different bilingual communities.

Conclusion

This study investigated the relationship between self-assessment of language proficiency and objective measures of lexical diversity and syntactic complexity in bilingual speakers of Italian in Croatia. Our primary aim was to investigate how these measures contribute to a comprehensive language profile of bilingual speakers obtained through their self-assessments. Exploring the informativeness of measures of lexical diversity and syntactic complexity is important as they can play an important role in the assessment of language proficiency. Firstly, these measures have proven to be ecologically valid as they capture language features in contexts that represent everyday language use, as opposed to the more formal settings of traditional language testing. Some speakers may see formal setting of proficiency tests stressful and unnatural, which may affect their performance. Using more ecological assessment offers them a better setting to showcase their abilities. Furthermore, there are no proficiency tests for all languages, which makes the measurement of lexical diversity and syntactic complexity a versatile tool that can be used in different linguistic contexts. By using an ecological assessment paradigm that relies on language samples rather than standardized tests, our research contributes to the ongoing discourse on the reliability of ecological measures of syntactic complexity and lexical diversity. The correlations found between these measures and different domains of self-assessment suggest that bilingual Italian speakers in Croatia are able to objectively assess their language proficiency. This is in line with the results of previous studies (Lu, 2012; Ortega, 2003; Read, 2000; Treffers-Daller, 2013), which confirm the credibility of self-assessment as a valuable tool for assessing the language proficiency of bilingual speakers.

This study also investigated language variety-related differences in the relationship between self-assessment and objective measures. We observed a higher number of correlations for dialect compared to standard Italian, which is probably due to the fact that the analyzed speech samples of spontaneous conversation were conducted in dialect, reflecting its nature as a spoken language variety. This emphasizes the complex interplay between literacy practices and dialect use. Future studies could investigate whether self-assessment in the standard variety correlates better with other objective measures of language proficiency, such as language tests. However, it is important to note that the dialect serves as a community language, and not all bilingual speakers in our sample attended Italian schools, which means that they may have limited access to the standard variety. Therefore, language assessment in dialect can provide more information about the language proficiency of bilingual speakers who live in such communities. For this purpose, both subjective forms of assessment such as self-assessment in dialect and objective assessments that are suitable for assessing dialect proficiency can be used. In contrast to language tests used to measure language proficiency, the indices used in language sample analysis are not limited to a specific language variety.

This study of Italophone bilingual speakers in Croatia sheds light on an understudied bilingual community. Findings might extend beyond Istria, offering insights applicable to similar bilingual communities facing diglossia challenges.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by The Ethics Committee of the Faculty of Education and Rehabilitation Sciences, University of Zagreb. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

GH: Conceptualization, Data curation, Formal analysis, Funding acquisition, Validation, Writing – original draft, Writing – review & editing. SK: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Validation. NP: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Data curation.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. GH, SK, and NP were funded by the Croatian Science Foundation (grant no. UIP-2017-05-6603, Multilevel Approach to Multilevel approach to spoken discourse in language development – MultiDis project). SK was funded by the Norwegian Research Council (grant no. 316103, Multilingual Minds: grammar interaction in multilingual acquisition project). The research was supported by the Croatian Science Foundation under the project Multilevel Approach to Multilevel approach to spoken discourse in language development – MultiDis (UIP-2017-05-6603).

Acknowledgments

The authors would like to thank the participants of this study, the investigators, and transcribers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., and Bialystok, E. (2018). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Resour.* 50, 250–263. doi: 10.3758/s13428-017-0867-9

Andrade, H., and Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Pract.* 48, 12–19. doi: 10.1080/00405840802577544

Andreou, M., and Tsimpli, I. M. (2020) in "Bilingualism, biliteracy and syntactic complexity: The role of crosslinguistic influence and cognitive skills" in language acquisition, processing and bilingualism: Selected papers from the romance turn VII, 132. Eds. A. Cardinaletti, C. Branchini, G. Giusti and F. Volpato (Newcastle upon Tyne: Cambridge Scholars Publishing), 132–159.

Bachman, L. F. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.

Bachman, L. F., and Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Lang. Test.* 6, 14–29. doi: 10.1177/026553228900600104

Beaudrie, S., and Ducar, C. (2005). Beginning level university heritage programs: creating a space for all heritage language learners. *Heritage Lang. J.* 3, 1–26. doi: 10.46538/hlj.3.1.1

Bedore, L. M., Peña, E. D., Gillam, R. B., and Ho, T. H. (2010). Language sample measures and language ability in Spanish-English bilingual kindergarteners. *J. Commun. Disord.* 43, 498–510. doi: 10.1016/j.jcomdis.2010.05.002

Biber, D. (1988). Variation across speech and writing. Cambridge: Cambridge University Press.

Birjandi, P., and Bolghari, M. S. (2015). The relationship between the accuracy of self- and peer-assessment of Iranian intermediate EFL learners and their learning styles. *Theory Pract. Lang. Stud.* 5, 996–1006. doi: 10.17507/tpls.0505.15

Bishop, D., and Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narrative: evidence for specific language impairment. *Br. J. Dev. Psychol.* 23, 25–46. doi: 10.1348/026151004X20685

Blagoni, R., Poropat Jeletić, N., and Blecich, K. (2016). "The Italophone reefs in the Croatophone Sea" in Bilingual landscape of the contemporary world. Eds. S. Grucza, M. Olpińska-Szkiełko and P. Romanowski (Wien: Peter Lang Verlag), 11–36.

Blanche, P., and Merino, B. J. (1989). Self-assessment of foreign-language skills: implications for teachers and researchers. *Lang. Learn.* 39, 313–338. doi: 10.1111/j.1467-1770.1989.tb00595.x

Blanchet, P., and Schiffman, H. (2004). Revisiting the sociolinguistics of Occitan: a presentation. *Int. J. Sociol. Lang.* 2004, 3–24. doi: 10.1515/ijsl.2004.042

Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge: what can heritage language learners contribute? *Stud. Second. Lang. Acquis.* 33, 247–271. doi: 10.1017/S0272263110000756

Brantmeier, C., Vanderplank, R., and Strube, M. (2012). What about me? *System* 40, 144–160. doi: 10.1016/j.system.2012.01.003

Brehmer, B. (2021). Maintenance of Russian as a heritage language in Germany: a longitudinal approach. *Russ. J. Linguist.* 25, 855–885. doi: 10.22363/2687-0088-2021-25-4-855-885

Brown, G. T. L., and Harris, L. R. (2013). "Student Self-Assessment" in The SAGE handbook of research on classroom assessment. ed. J. H. McMillan (Los Angeles/London: SAGE), 367–393.

Caissie, R., Dawe, A. L., Donovan, C., Brooks, H., and MacDonald, S. M. (1998). Conversational performance of adults with hearing loss. *J. Acad. Rehabil. Audiol.* 36, 45–67.

Clyne, M. (1997). Language policy in Australia – achievements, disappointments, prospects. *J. Intercult. Stud.* 18, 63–71. doi: 10.1080/07256868.1997.9963442

Covington, M. A., and McFall, J. D. (2010). Cutting the Gordian knot: the moving-average type-token ratio (MATTR). *J. Quantitative Linguist.* 17, 94–100. doi: 10.1080/09296171003643098

Creaghe, N. V. (2019). Symbolic play and language acquisition: the dynamics of infant-caregiver communication during symbolic play. [dissertation/master's thesis]. Canberra: The Australian National University. Available at: https://openresearch-repository.anu.edu.au/handle/1885/201973 (Accessed October 15, 2023).

Davidson, D. E., and Lekic, M. D. (2013). The heritage and non-heritage learner in the overseas immersion context: comparing learning outcomes and target-language utilization in the Russian flagship. *Heritage Lang. J.* 10, 226–252. doi: 10.46538/hlj.10.2.6

De Angelis, G. (2007). Third or additional language acquisition. Clevedon: Multilingual Matters.

Delgado, P., Guerrero, G., Goggin, J. P., and Ellis, B. B. (1999). Self-assessment of linguistic skills by bilingual Hispanics. *Hisp. J. Behav. Sci.* 21, 31–46. doi: 10.1177/0739986399211003

DeVellis, R. F. (2003). Scale development: Theory and applications. Thousand Oaks: SAGE Publications.

Durrant, P., Brenchley, M., and McCallum, L. (2021). Understanding development and proficiency in writing: Quantitative Corpus linguistic approaches. Cmabridge: Cambridge University Press.

Ebert, K. D. (2020). Language sample analysis with bilingual children. *Lang. Disord.* 40, 182–201. doi: 10.1097/TLD.0000000000000209

Fagan, J., and Iglesias, A. (2000). The relationship between fathers' and children's communication skills and children's behavior problems: a study of head start children. *Early Educ. Dev.* 11, 307–320. doi: 10.1207/s15566935eed1103_5

Fergadiotis, G., Wright, H. H., and Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: assessing length effects. *J. Speech Lang. Hear. Res.* 58, 1–13.

Flege, J. E., Yeni-Komishian, G., and Liu, S. (1999). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Biling. Lang. Congn.* 3, 131–150.

Frizelle, P., Thompson, P. A., McDonald, D., and Bishop, D. V. M. (2018). Growth in syntactic complexity between four years and adulthood: evidence from a narrative task. *J. Child Lang.* 45, 1174–1197. doi: 10.1017/S0305000918000144

Giuricin, E., and Giuricin, L. (2008). La comunità nazionale italiana. Storia e istituzioni degli Italiani dell'Istria, Fiume e Dalmazia (1944–2006). Rovinj/Rovigno: Centro di ricerche storiche di Rovigno.

Gutierrez-Clellen, V. F., and Hofstetter, R. (1994). Syntactic complexity in Spanish narratives: a developmental study. *J. Speech Hear. Res.* 37, 645–654. doi: 10.1044/jshr.3703.645

Hayakawa, S., Chung-Fat-Yim, A., and Marian, V. (2022). Predictors of language proficiency and cultural identification in heritage bilinguals. *Front. Commun.* 7, 1–24. doi: 10.3389/fcomm.2022.994709

He, A. W. (2010). The heart of heritage: sociocultural dimensions of heritage language learning. *Annu. Rev. Appl. Linguist.* 30, 66–82. doi: 10.1017/S0267190510000073

Heilmann, J., Miller, J. F., Nockerts, A., and Dunaway, C. (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *Am. J. Speech Lang. Pathol.* 19, 154–166. doi: 10.1044/1058-0360(2009/08-0024)

Hoff, E., Core, C., Place, S., and Rumiche, R. (2012). Dual language exposure and early bilingual development. *J. Child Lang.* 39, 1–27. doi: 10.1017/S0305000910000759

Hržica, G., and Roch, M. (2021). "Lexical diversity in bilingual speakers of Croatian and Italian" in LITMUS in action: Crosscomparison studies across Europe. Eds. S. Armon-Lotem and K. Grohmann, vol. *29* (Amsterdam: John Benjamins).

Iannaccaro, G., Dell'Aquila, V., and Andrione, E. (Eds.) (2003). Une Vallée d'Aoste bilingue dans une Europe plurilingue/Una Valle d'Aosta bilingue in un'Europa plurilingue. Aosta: Fondation Emile Chanoux.

IBM Corp (2011). IBM SPSS statistics for windows: Version 20.0. Armonk, NY: IBM Corp.

Ikeguchi, C. B. (1996). Self-assessment and ESL competence of Japanese returnees. Research/Technical Report. Available at: http://eric.ed.gov/PDFS/ED399798.pdf (Accessed October 15, 2023).

Jarvis, S. (2002). Short texts, best–fitting curves and new measures of lexical diversity. *Lang. Test.* 19, 57–84. doi: 10.1191/0265532202lt220oa

Kang, H., and Kim, I. (2012). Perceived and actual competence and ethnic identity in heritage language learning: a case of Korean-American college students. *Int. J. Biling. Educ. Biling.* 15, 279–294. doi: 10.1080/13670050.2011.626846

Kang, H., and Kim, I. (2016). Heritage language self-assessment: the role of cultural identity and language domain. *Kor. J. Appl. Linguist.* 32, 27–50. doi: 10.17154/kjal.2016.9.32.3.27

Kelly, L., Nixon, E., and Quigley, J. (2022). It's your turn: the dynamics of conversational turn-taking in father-child and mother-child interaction. *Lang. Develop. Res.* 2, 37–68. doi: 10.34842/840g-2297

Košutar, S., Kramarić, M., and Hržica, G. (2022). The relationship between narrative microstructure and macrostructure: differences between six-and eight-year-olds. *Psychol. Lang. Commun.* 26, 126–153. doi: 10.2478/plc-2022-0007

Krausert, S. R. (1991). Determining the usefulness of self-assessment of foreign language skills: post-secondary ESL students' placement contribution. *Opto Electronic Engin.* 4, 85–100.

Langan, A. M., Shuker, D. M., Cullen, W. R., Penney, D., Preziosi, R. F., and Wheater, C. P. (2008). Relationships between student characteristics and self-, peer and tutor evaluations of oral presentations. *Assess. Eval. High. Educ.* 33, 179–190. doi: 10.1080/02602930701292498

Lanthaler, F. (2001). "Zwischenregister der deutschen Sprache in Südtirol" in Die deutsche Sprache in Südtirol. Einheitssprache und regionale Vielfalt. Eds. K. Egger and F. Lanthaler (Wien/Bozen: Folio Verlag), 137–152.

Li, M., and Zhang, X. (2021). A meta-analysis of self-assessment and language performance in language testing and assessment. *Lang. Test.* 38, 189–218. doi: 10.1177/0265532220932481

Loban, W. (1976). Language development: Kindergarten through grade twelve. NCTE committee on research report no. 18. Urbana: National Council of Teachers of English.

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *Mod. Lang. J.* 96, 190–208. doi: 10.1111/j.1540-4781.2011.01232_1.x

Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574

Ma, W., and Winke, P. (2019). Self-assessment: how reliable is it in assessing oral proficiency over time? *Foreign Lang. Ann.* 52, 66–86. doi: 10.1111/flan.12379

Mac Whinney, B. (2000). The CHILDES project: Tools for analyzing talk: Transcription format and programs. *3rd* Edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Macbeth, A., Atagi, N., Montag, J. L., Bruni, M. R., and Chiarello, C. (2022). Assessing language background and experiences among heritage bilinguals. *Front. Psychol.* 13, 1–17. doi: 10.3389/fpsyg.2022.993669

MacIntyre, P. D., Noels, K. A., and Clément, R. (1997). Biases in self-ratings of second language proficiency: the role of language anxiety. *Lang. Learn.* 47, 265–287. doi: 10.1111/0023-8333.81997008

Mäkinen, L., Loukusa, S., Nieminen, L. P., Leinonen, E., and Kunnari, S. (2014). The development of narrative productivity, syntactic complexity, referential cohesion and event content in four-to eight-year-old Finnish children. *First Lang.* 34, 24–42. doi: 10.1177/0142723713511000

Malvern, D., and Richards, B. J. (1997). "A new measure of lexical diversity", in evolving models of language. Clevedon: Multilingual Matters, 58–71.

Malvern, D., Richards, B. J., Chipere, N., and Durán, P. (2004). Lexical diversity and language development: Quantification and assessment. Basingstoke, UK: Palgrave Macmillan.

Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Resour.* 50, 940–967. doi: 10.1044/1092-4388(2007/067)

McCarthy, P. M., and Jarvis, S. (2010). MTLD, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behav. Res. Methods* 42, 381–392. doi: 10.3758/BRM.42.2.381

Milani Kruljac, N. (2001). "Situazione linguistica" in La comunità rimasta. Eds. F. Radin and G. Radossi (Zagreb: Garmond), 239–250.

Milani Kruljac, N. (2003). L'italiano fra i giovani dell'Istroquarnerino. Pola-Fiume: Pietas Iulia-Edit.

Miller, J. F. (1981). Assessing language production in children. Baltimore, MD: University Park Press.

Miller, H., and Miller, K. (1996). Language policy and identity: the case of Catalonia. *Int. Stud. Sociol. Educ.* 6, 113–128.

Mitrofanova, N., Rodina, Y., Urek, O., and Westergaard, M. (2018). Bilinguals' sensitivity to grammatical gender cues in Russian: the role of cumulative input, proficiency, and dominance. *Front. Psychol.* 9:1894. doi: 10.3389/fpsyg.2018.01894

Montrul, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. *Stud. Second. Lang. Acquis.* 33, 155–161. doi: 10.1017/S0272263110000719

Montrul, S. (2015). The Acquisition of Heritage Languages. Cambridge: Cambridge University Press.

Montrul, S., and Polinsky, M. (2021). The Cambridge handbook of heritage languages and linguistics. Cambridge: Cambridge University Press.

Ninio, A., and Snow, C. (1999). "The development of pragmatics: learning to use language appropriately" in Handbook of child language acquisition. Eds. W. C. Ritchie and T. K. Bhatia (Leiden: Brill Academic Press), 347–383.

Nitta, R., and Nakatsuhara, F. (2014). A multifaceted approach to investigating pre-task planning effects on paired oral test performance. *Lang. Test.* 31, 147–175. doi: 10.1177/0265532213514401

Norris, J. M., and Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Appl. Linguis.* 30, 555–578. doi: 10.1093/applin/amp044

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: a research synthesis of college-level L2 writing. *Appl. Linguis.* 24, 492–518. doi: 10.1093/applin/24.4.492

Oscarson, M. (1989). Self-assessment of language proficiency: rationale and applications. *Lang. Test.* 6, 1–13. doi: 10.1177/026553228900600103

Panadero, E., and Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment Educ.* 21, 133–148. doi: 10.1080/0969594X.2013.877872

Peltonen, P. (2021). Connections between measured and assessed fluency in L2 peer interaction: a problem-solving perspective. *IRAL* 60, 983–1011. doi: 10.1515/iral-2020-0030

Polinsky, M. (2014). "Heritage languages and their speakers: Looking ahead", in innovative approaches to heritage languages: From research to practice. Washington, D.C.: Georgetown University Press, 1–24.

Polinsky, M. (2018). Heritage languages and their speakers. Cambridge: Cambridge University Press.

Poropat Jeletić, N., Moscarda Mirković, E., and Hržica, G. (2024). BilingBank the Corpus of spoken Istrovenetian/Fiuman and Croatian (C-ORAL-IC).

Read, J. (2000). Assessing vocabulary. Cambridge: Cambridge University Press.

Regula, C. (2008). Sprachgebrauch im Alltag–Die Verwendung des Rätoromanischen in verschiedenen Domänen: Wechselwirkungen und Einflussfaktoren. Chur: Institut für Kulturforschung Graubünden.

Roland, D., Dick, F., and Elman, J. (2007). Frequency of basic English grammatical structures: a corpus analysis. *J. Mem. Lang.* 57, 348–379. doi: 10.1016/j.jml.2007.03.002

Romano, F., and Guijarro-Fuentes, P. (2023). Task effects and the yes-bias in heritage language bilingualism. *Int. J. Biling. Educ. Biling.* 27, 389–409. doi: 10.1080/13670050.2023.2206949

Romano, F., and Sorace, A. (2023). "Conclusions" in Studies in Italian as a heritage language. ed. F. Romano (Berlin: Mouton de Gruyter).

Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Lang. Test.* 15, 1–20. doi: 10.1177/026553229801500101

Ross, S. (2006). The reliability, validity, and utility of self-assessment. *Pract. Assess. Res. Eval.* 11, 1–13.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: romance languages as heritage languages. *Int. J. Biling.* 13, 155–163. doi: 10.1177/1367006909339814

Rothman, J., Bayram, F., DeLuca, V., González Alonso, J., Kubota, M., and Puig Mayenco, E. (2023). "Defining bilingualism as a continuum: Some tools and consequences for the study of bilingual mind and brain effects" in Understanding Language and Cognition through Bilingualism: In honor of Ellen Bialystok (Studies in Bilingualism). Eds. G. Luk, J. A. E. Anderson and J. G. Grundy, vol. *64* (Amsterdam: John Benjamins), 38–67.

Selinker, L., and Baumgartner Cohen, B. (1995). Multiple language acquisition: 'damn it, why Can't I keep these two languages apart. *Lang. Cult. Curric.* 8, 115–121. doi: 10.1080/07908319509525195

Talamas, A., Kroll, J. F., and Dufour, R. (1999). From form to meaning: stages in the acquisition of second-language vocabulary. *Biling. Lang. Congn.* 2, 45–58. doi: 10.1017/S1366728999000140

Templin, M. (1957). Certain language skills in children. Minneapolis: University of Minnesota Press.

Thompson, G. L. (2015). Understanding the heritage language student: proficiency and placement. *J. Hisp. High. Educ.* 14, 82–96. doi: 10.1177/1538192714551277

Titus, J. (2012). Creating an online assessment test for heritage learners of Russian. *Heritage Lang. J.* 9, 222–244. doi: 10.46538/hlj.9.2.6

Treffers-Daller, J. (2013). "Measuring lexical diversity among L2 learners of French: An exploration of the validity of D, MTLD and HD-D as measures of language ability" in vocabulary knowledge: Human rating and automated measures. Eds. S. Jarvis and M. Daller (Amsterdam: Benjamins), 79–104.

Treffers-Daller, J. (2015). "Language dominance: the construct, its measurement, and operationalization" in Language dominance in bilinguals: issues of measurement and operationalization. Eds. C. Silva-Corvalán and J. Treffers-Daller (Cambridge: Cambridge University Press), 235–265.

Treffers-Daller, J., and Korybski, T. (2015). "Using lexical diversity measures to operationalize language dominance in bilinguals" in Language dominance in bilinguals: Issues of measurement and operationalization. Eds. C. Silva-Corvalán and J. Treffers-Daller (Cambridge: Cambridge University Press), 106–133.

Treffers-Daller, J., Parslow, P., and Williams, S. (2018). Back to basics: how measures of lexical diversity can help discriminate between CEFR levels. *Appl. Linguis.* 39, amw009–amw327. doi: 10.1093/applin/amw009

Unsworth, S., Brouwer, S., de Bree, E., and Verhagen, J. (2019). Predicting bilingual preschoolers' patterns of language development: degree of non-native input matters. *Appl. Psycholinguist.* 40, 1189–1219. doi: 10.1017/S0142716419000225

Van der Veen, C., Michaels, S., Dobber, M., van Kruistum, C., and van Oers, B. (2021). Design, implementation and evaluation of dialogic classroom talk in early childhood education. *Learn. Cult. Soc. Interact.* 29, 100515–100514. doi: 10.1016/j.lcsi.2021.100515

Van Os, M., de Jong, N. H., and Bosker, H. R. (2020). Fluency in dialogue: turn-taking behavior shapes perceived fluency in native and nonnative speech. *Lang. Learn.* 70, 1183–1217. doi: 10.1111/lang.12416

Williams, S., and Hammarberg, B. (1998). Language switches in L3 production: implications for a polyglot speaking model. *Appl. Linguis.* 19, 295–333. doi: 10.1093/applin/19.3.295

Zell, E., and Krizan, Z. (2014). Do people have insight into their abilities? A Metasynthesis. *Perspect. Psychol. Sci.* 9, 111–125. doi: 10.1177/1745691613518075