ISSN 2385-4138 (digital)
https://doi.org/10.5565/rev/isogloss.302

LexSIC: A quick vocabulary test for Sicilian

Tanja Kupisch
University of Konstanz / UiT The Arctic University of Norway
tanja.kupisch@uni-konstanz.de

Sebastiano Arona
University of Konstanz
sebastiano.arona@uni-konstanz.de

Alexandra Besler
University of Konstanz
alexandra.besler@uni-konstanz.de

Silvio Cruschina
University of Helsinki
silvio.cruschina@helsinki.fi

Maria Ferin
University of Konstanz
maria.ferin@uni-konstanz.de

Henrik Gyllstad
University of Lund
henrik.gyllstad@englund.lu.se

Ilaria Venagli
University of Konstanz
ilaria.venagli@uni-konstanz.de

Received: 23-01-2023
Accepted: 05-05-2023
Published: 17-05-2023

How to cite: Kupisch, Tanja, Arona, Sebastiano, Besler, Alexandra, Cruschina, Silvio, Ferin, Maria, Gyllstad, Henrik, Venagli, Ilaria. 2023. LexSIC: A quick vocabulary test for Sicilian. *Isogloss. Open Journal of Romance Linguistics* 9(1)/6, 1-24.
DOI: https://doi.org/10.5565/rev/isogloss.302

Abstract

It has been increasingly acknowledged that regional languages and dialects bear similarities with heritage languages, and that some heritage speakers are bilectal with two closely related minority languages or dialects. However, assessing the knowledge of non-standard varieties proves difficult due to the lack of assessment materials. Our contribution documents the creation and validation of LexSIC, a yes/no vocabulary task, a placement test for speakers with receptive and/or productive skills in Sicilian. The materials are freely available. The vocabulary items were validated in a survey with 100 speakers of Italian, who had varying levels of proficiency in Sicilian, including bilectal heritage speakers in Germany. We have relied on Classical Test Theory (CTT) and Item Response Theory (IRT) in validating test items and we demonstrate and discuss correlations with self-reported measures of proficiency, language use, age and education. A comparison of vocabulary proficiency in bilectals with Italian and Sicilian proficiency further shows that the two are highly correlated, which underlines a cumulative enhancement of vocabulary proficiency in bilectals. We discuss the challenges in creating assessment materials in minority languages and dialects.

Keywords: Sicilian, bilectalism, heritage speakers, vocabulary, assessment.

1. Introduction

This paper reports on the creation of LexSIC, a vocabulary assessment task for Sicilian. Sicilian is a regional language in Italy which has several million speakers in the world (Berruto 2018; Eberhard et al. 2022). For political and sociolinguistic reasons, similar to many other varieties spoken in Italy (e.g., Ligurian, Venetan), Sicilian is traditionally referred to as an Italian or Italo-Romance "dialect" to mark its sociolinguistic subalternity to the standard language. This tradition is fostered in the linguistic literature, and it is also cultivated by many of the speakers themselves (Grassi et al. 1997; Loporcaro 2009; see Besler et al. in prep on Sicilians' attitudes towards Sicilian). However, these so-called "dialects" *of Italy*, including Sicilian, are not dialects *of Italian*, originating from the standard variety, but rather fully-fledged Romance languages spoken in Italy that have independently evolved from Latin. In this sense, they are sister languages to standard Italian. These languages are nowadays mostly used as spoken languages and are often bound to specific and limited contexts. By contrast, some modern varieties of Italian have evolved from standard Italian under the influence of the regional languages spoken in those areas, and they are generally called "regional varieties" of Italian. Herein, we take Sicilian to be a language as it represents an independent linguistic system in opposition to Italian. We will use the terms "bilectalism" and "bilectals in Sicily" to refer to the linguistic situation of (most) Sicilians who speak Sicilian and Standard Italian.

We wish to underline that the situation of Sicilian-Italian speakers is very similar to that of bilingualism, more specifically, *bilingualism with diglossia* (Ferguson 1964; Berruto 1987). This similarity further implies that "heritage bilingualism" (the acquisition of an immigrant minority language) and "bilectal

bilingualism" represent similar scenarios, because in both cases bilingual speakers acquire and use a minority language primarily in the home, although one difference is that regional languages and dialects are often autochthonous and the two varieties of bilectals tend to be typologically close.

Bilectal Italians are not only found in Italy, however. Italians have been very much on the move, at least since the unification of Italy in the 18th century. People of Italian descent, about 80 million, live in all parts of the world, including, e.g., significant minorities in South America, the US, Canada, Australia, Belgium, and Germany (see, e.g., Haller 1993; Rubino 2004; Andriani et al. 2022 for relevant linguistic studies). Thus, in addition to what we may refer to as "homeland bilectals", there are also "heritage bilinguals" in other parts of the world.

However, most studies on Italian heritage speakers have focused on their proficiency in (*Standard) Italian*, and considering the history of Italian migration, this is somewhat concerning. Many Italians have left their home country during a time when Standard Italian was spoken by a small minority of people inhabiting the geographical area that makes up today's Italy. By the time of the Unification of Italy in 1861 only 3% of the population spoke Standard Italian (Tosi 2001; De Mauro 2005). Speaking a local language was the norm, and contact with Standard Italian was the exception, as this was the time before TV and obligatory schooling (De Renzo 2008). Arrived in the diaspora, immigrants faced the challenge of learning a new language, while passing on their native language to their children —the local language in the place where they grew up, i.e., *not* Standard Italian. This does not mean to imply that today's "Italian" HS do not know Standard Italian. They do know Standard Italian for various reasons: First, also in the diaspora Italians had to communicate with each other, but since they did not always speak the exact same variety, they often resorted to one that was felt to be more easily mutually understandable, e.g., an Italianized version of their own variety. Second, a more Italian-like variety was considered to be more beneficial for the next generation, because many regional languages were stigmatized. Third, as a response to the increasingly positive attitude towards multilingualism in the Western World[1], the awareness that multilingualism might come with benefits for cognitive development (Bialystok et al. 2004) and for the society at large, there is growing willingness to help minorities preserve their ethnic cultures, languages and identities. It is increasingly common for HS to benefit from language classes or even education in a minority language as the language of instruction, organized by local culture institutes, embassies or local bilingual schools. Crucially, however, these language classes will be geared towards acquiring the standard language. Thus, our point is not that speakers with Italian heritage have no proficiency in Italian. Our point is that they might have *additional proficiencies* in another regional language/ dialect, such as Sicilian, which needs to be assessed if we want to understand heritage bilingualism (see Leivada et al. 2017 for a similar reasoning for Cypriot and Standard Modern Greek bilectals).

In linguistic research, including first, second, and third language acquisition research, Italians are typically taken to be speakers of Italian "only", plus the societal majority language and potential foreign languages, while proficiency in additional dialects or regional languages is not mentioned. Not only is this methodologically

[1] For example, the Council of Europe has committed to the protection and promotion of multilingualism (Council of Europe 1992).

unsound because these speakers are compared to Italians in Italy, who are continuously exposed to Standard Italian, including in formal contexts; but it is also a missed opportunity because there might be transfer effects into later acquired languages that cannot be predicted based on knowledge of Italian alone, as they result from additional knowledge of a related language/ dialect instead.[2] However, speculating about potential transfer from such an additional linguistic source is problematic without any assessment of the acquired knowledge in this source language/ dialect: What has not been acquired, cannot be transferred. Self-reported proficiency is a start, but generally considered to be insufficient and unreliable. This was the motivation for creating LexSIC.

The creation of a (quick) assessment tool for language proficiency is nothing new. There are indeed many well-documented examples, e.g., LexTALE, LexTALE_Fr, and LexITA (e.g., Lemhöfer & Broersma 2012; Brysbaert 2013; Amenta et al. 2021), all lexical decision tasks based on a choice between words and non-words. However, creating a task for a regional minority language or dialect comes with some additional challenges. First and foremost, regional languages/ dialects are a moving target because they are typically not standardized, even if there may be one variety that is officially or unofficially considered to be more neutral (Loporcaro 2009). Sicilian is a case in point. It may be considered "uniform" with respect to the written language, used in literary production (cf. the variety of *siciliano illustre* used by the *Sicilian school*), but its uniformity has been disputed for the spoken language, which is indeed subject to diatopic variation and lacks a standard variety (Coluccia 1974; Ruffino 1991, 2001; see also Matranga & Sottile 2013; Cruschina 2020 and references therein on the current linguistic situation of Sicilian and dialectal classification). Thus, we can expect there to be more variation, which language users are aware of and which they accept. We can further expect that regional languages and dialects are more susceptible to language change due to the absence of continuous norming processes, as we typically see them for national standard languages, which are used and taught in educational settings. It also means that even with the best of all tests, which take variation into account, we cannot expect potential test takers, including highly proficient native speakers, to score at ceiling. Second, and relatedly, regional languages and dialects do not always have a written standard, and even if they do, speakers may not be familiar with it. Consequently, it will not be suitable to present stimuli in an exclusively written mode. Third, almost every speaker of a dialect/ regional minority language today is bilingual because their varieties coexist in a diglossic situation with the standard language (Berruto 1987, 1989). As to speakers in the diaspora, they speak another, often typologically distant, majority language. This means that there cannot be a monolingual control group, as is customary when creating new assessment tools. Fourth and finally, there are the practical issues of finding enough speakers to pilot the task. For example, LexITA was piloted with 199 speakers of Italian (58 L1 and 141 L2) but there are 85 million Italian speakers in the world, and Italian can be learnt in foreign language classrooms. This is a not a luxury we have with Sicilian. Sicilian is not a national language and it is not taught as a foreign language either. Who comes to Sicily as an immigrant is more likely to find themselves

[2] An alternative scenario is that the parents/grandparents of the HS are bilectal and spoke Italian in the home but with transfer from the dialect. In this scenario, the HS themselves may not be bilectal, but their Italian might have Sicilian traits (as is the case for regional Italian), which can be transferred to the majority language.

in an Italian classroom. The validation and interpretation of data for dialect assessment must acknowledge these challenges.

2. Background

2.1. Sicilian

Sicilian (*sicilianu*) belongs to the Extreme Southern Italian language group (*dialetti italiani meridionali estremi*) spoken in the South of Italy, which covers the Island of Sicily as well as the Southern parts of Calabria and the Salento (Pellegrini 1977). As mentioned above, the Ethnologue (Eberhard et al. 2022) indicates a number of 4,7 million speakers of Sicilian worldwide. Following Berruto (2018: 202), based on ISTAT (2012), 72% of Sicily's population (about 3,5 million) uses Sicilian actively. Sicilian is also spoken by a large number of people in the diaspora, including Australia (Rubino 2014), the US (Haller 1993, 1997), South America (Andriani et al. 2022) and Germany. According to Schmid (2014), in 2008 35% of the Italians living in Baden-Württemberg (German federal state) had Sicilian roots (roughly 57,000). Following Pellegrini (1977), Sicilian can be divided into seven dialects, including Western Sicilian (including Palermo, Trapani), Central, Southeast, East, Messinese, Isole Eolie, Pantesco. Southern Calabrian and Salentino can also be considered dialects of Sicilian.

The Ethnologue (Eberhard et al. 2022) classifies Sicilian as a separate language because it is distinct enough from Standard Italian; Sicilian is also recognized as a minority language by the UNESCO. The Ethnologue further considers Sicilian a "large" language (more than 1 million speakers) and a "stable" language, which means, according to their description, that although the language is not being sustained by formal institutions, it is still the norm in the home and community that many children learn and use the language. However, the Italian authorities have not recognized Sicilian as a language, similar to other regional languages in Italy (Moseley 2010: 39), and, as discussed, it is common to speak of Sicilian in terms of a "dialect", including amongst Sicilians themselves. This is why relevant surveys in Italy typically report "dialect use".

In the South of Italy, the use of regional languages and dialects is generally more widespread than in the Centre and in the North West. Sicily, together with Calabria, Campania and Veneto is amongst the regions where use of the local varieties is most widespread, with 25.5% of the population claiming that they use mostly or exclusively dialect, 46.2% using both dialect and Italian and only 26.2% using only Italian (ISTAT 2007). The use of regional languages/ dialect is determined by age and education. While only 20.3% of people with a university degree report dialect use, the percentage amounts to 72.4% amongst people with only a primary school certificate (Berruto 2018: 504, based on ISTAT 2012). While regional languages/ dialects in Italy have lost speakers over time, the attitudes towards them have changed for the better: they are no longer stigmatized and considered to be a marker of a lower educational status, at least not generally. However, as Ruffino (2006) has pointed out, the revaluation of dialects has not happened to the same degree in all regions, and negative attitudes, especially towards the Southern varieties, which are overall more widespread, may persist.

2.2 Vocabulary knowledge tests

Generally speaking, vocabulary knowledge is conventionally modelled as having three dimensions: size, depth and fluency (Gyllstad 2013). For the present purposes, the dimension of size is the most relevant (see e.g., Schmitt, 2014 for accounts of depth and fluency). Vocabulary size, also referred to as *breadth*, refers to the number of words for which a language user possesses at least a basic form-meaning mapping, i.e., all word forms that the language user can map onto a conventionalized meaning. Meara (1996) considers *size* the most basic dimension, and it is considered to be a good indicator of general language proficiency (Anderson & Freebody 1983; Gathercole & Baddeley 1989; Nation 1993; Bates & Goodman 1999; Hilton 2008; Staehr 2008; Laufer & Ravenhorst-Kalovski 2010; Jeon & Yamashita 2014; Zhang & Zhang 2020). Desirable qualities in size tests that can be used for research purposes are that it measures lexical knowledge objectively, reliably, and rapidly, so that it can easily be integrated in a study. A receptive task is ideal because it provides us with the possibility to target lower proficiency level users of a language, which would be impossible with a speaking or writing task.

The particular vocabulary knowledge test format relevant to the present paper is called a "yes/no" test (Meara & Buxton 1987). The yes/no format is commonly used in vocabulary size tests. Although the purpose of our test is not technically to yield a size estimate per se, since this would require a systematic stratified random sample from a defined population of Sicilian words, it still shares many characteristics with bona fide size tests. It presents a list of target words and typically asks whether the meaning of each word is known, or whether the word exists. It technically corresponds to a self-reported form recognition format. Test-takers indicate whether the presented words exist and/or are known. Depending on the instruction wording, it could be argued to be a measure of meaning recall, which presupposes that the word form is recognized. While there is no verification whether the words *indicated* as "existing" or "known" *are* actually known by the test-taker, yes/no tests typically include pseudo-words, whose function is to counter overestimation of test-taker scores. This is done by score-wise penalizing a test-taker when answering "yes" to a pseudo-word. The yes/no test format has played a significant role in the DIALANG project, whose goal was to provide an online tool for language assessment in 14 European languages, including Italian (Alderson 2005; Alderson et al. 2005). The DIALANG test suite contains an initial placement test, used to determine which difficulty level is appropriate for the continued testing. These placement tests in all languages consist of 75 items, with 50 real and 25 pseudo-verbs in the infinitive form, sampled from dictionaries. The use of verbs was argued to be an effective means of accessing a relatively large range of a given language's vocabulary with only few items (Alderson 2005: 80). Our test will largely be modeled after the DIALANG with some modifications outlined below.

In what follows, we describe the development and validation of LexSIC in detail. The measures of vocabulary that have been developed and validated recently have often relied on Classical Test Theory (CTT) and Item Response Theory (IRT), a procedure we will follow here (see, e.g., Beglar 2010; Gyllstad et al. 2015, 2021; McLean et al. 2015). We first describe the initial version of the test and then how we assessed the items by means of CTT and IRT. In section 4, we present the correlations

between LexSIC and some variables for bilectal speakers in Sicily and Germany as well as two control groups.

3. Method

3.1 Creating the material for LexSIC

We followed the DIALANG procedure of selecting existing verbs and creating additional pseudo-verbs with the goal to arrive at a final test with 75 items (50 real + 25 pseudo-verbs). The reason for creating a task that resembles the DIALANG placement test (rather than LexITA) was that the DIALANG placement tests already exist for many other languages. This way, assessment of lexical proficiency in heritage bilectals in the diaspora can be done on the basis of comparable tests in three or more languages. Thus, the choice of using only verbs was to achieve the highest possible comparability with the DIALANG. The other reason was that nouns were felt to introduce more noise, as they vary more across regions.[3]

We preselected 80 real verbs on the basis of Sicilian dictionaries and Sicilian literature. We ensured that the majority of verbs looked phonologically and morphologically different from their Italian equivalents and that they were not archaic. We then created 48 pseudo-verbs, using three different strategies. The first was to add the verbal suffix *-ari* or *-iri* to existing noun stems (e.g., sic. *seggia* 'chair' > segg(i) *+-ari* → **seggiari*). The second was to add a Sicilian suffix to foreign verb stems, e.g., en. *type*, ge. *tippen* > *tip* + *-ari* → **tippari*. The third was to manipulate existing verbs phonologically (e.g., *criticari* → *critinari*). The pseudo-verbs had the phonotactic and morphological properties of Sicilian verbs. This preliminary list was given to five native speakers of Sicilian from different areas in Sicily (South, East, West and Centre). They were asked to rank the pseudo-verbs in terms of plausibility on a scale from 1 to 5. For the real words, we asked our informants whether they knew these verbs and, if yes, to indicate the estimated frequency on a 5-point scale. We discussed the results with our informants.

Based on this preliminary assessment, we reduced the number of items to a subset of 103 (see Appendix A1 for the final list of items). Amongst the pseudo-verbs we removed 10 items, making sure the remainder was distributed evenly over the five plausibility ranges. For the real words, one problem that became evident was that high perceived difficulty generally coincided with diatopic variation, as items that were unknown to several participants tended to be specific to particular areas, towns or even villages in Sicily. Thus, the elimination of items that could potentially favor speakers from particular regions, also led to the exclusion of some items at the higher level of the difficulty scale. Another problem that became apparent was that pronunciation could differ slightly across regions. However, creating separate tasks per area was neither a goal nor an option for us. We therefore decided to keep items with phonological variants but alleviate the problem by adding a passage in the instruction to remind participants of geographical variation in pronunciation and provide some examples (see Appendix A2). Another problem we had to address concerned the fact

[3] In this way, we deviate from LexITA which also included nouns and adjectives. We return to this point in the discussion.

that Sicilian is primarily a spoken language. Hence, presenting stimuli in a written mode would come with the risk that people might not be familiar with the chosen orthography and/or misinterpret it. We therefore decided to present the items in a dual mode. The remaining 103 items were consequently recorded by a native Sicilian speaker from Palermo and double-checked by another native speaker of Sicilian (there was disagreement on the stress pattern of one item, which had to be recorded again).

3.2 Participants

The main participants of our study were bilectals in Sicily and bilectals in the diaspora (bilectal HS). We further included a handful of speakers from regions which are part of the Extreme Southern language group outside of Sicily (Calabria and Puglia) as well as Northern Italians. We did not expect to find speakers who have grown up monolingually, because nowadays speakers of Sicilian have additional language experience in the majority language Italian. This explains the absence of a "monolectal" control group. Moreover, while other studies have piloted assessment tasks with L2 learners (e.g., Lemhöfer & Broersma 2012; Amenta et al. 2021; Salmela et al. 2021), it was unrealistic to find L2 learners of Sicilian (with L2 in the "classroom" sense), because Sicilian is not taught as a foreign language.[4] However, HS are comparable to L2 learners in the sense that they typically represent a large range of proficiencies. The two control groups were added to represent an even greater range of proficiencies. Overall, the sample included 100 participants (see Table 1).

The highest proficiency level of Sicilian was expected amongst speakers residing in Sicily, who represent our largest group (n = 56). The speakers from North Italy were included to represent a low proficiency level, and to identify verb items whose meaning can be guessed easily given knowledge of Italian. The speakers from South Italy outside of Sicily are expected to understand Sicilian to a good extent, though not necessarily each individual item, due to geographical distance and variation.

Table 1. Speaker Groups and biographical data. The scale for self-rated Sicilian proficiency ranged from 1-6; that for Sicilian use from 0-25 (0=no use of Sicilian, 25=exclusive use of Sicilian)

	N	**Mean age (*SD*; range)**	**Self-rated proficiency, mean (*SD*; range)**		**Sicilian use, mean (*SD*; range)**	
			productive	***receptive***	***home***	***outside home***
Bilectals in Sicily	56	34.8 (*11.31*; 18-69)	4.39 (*1.61*; 1-6)	5 (*1.16*; 1-6)	9.87 (5.*33*; 0-4)	5.82 (*4.59*; 0-28)
Bilectal HS	19	28.5 (*9.93*; 14-53)	3.05 (*2.04*; 1-6)	4.05 (*1.75*; 1-6)	8.41 (*7.77*; 0-4)	2.66 (*5.53*; 0-24.5)
Non-Sicilian Southern Italians	5	38.0 (*10.99*; 23-52)				
Northern Italians	10	31.3 (*10.89*; 24-62)				

[4] It is possible for someone to move to Sicily at a later age and acquire the dialect, but it is more likely that this person acquires Italian.

Detailed biographical information and self-rated proficiency in Sicilian was only collected from the two main groups of Italian-Sicilian bilectals. In terms of educational level, most bilectal HS had university (n = 6) or high school degrees (n = 8), while only 5 had middle school or lower degree. Amongst the Sicilian bilectals too, most had high school (n = 19) or university degrees (n = 33); only 4 had middle school[5] or lower degree. Data on self-rated proficiency ranged from 1 (min) to 6 (max) in both groups, both receptive and productive skills. That is, both groups used the entire scale when rating their own skills (see Table 1). Bilectals from Sicily reported higher productive and receptive skills, and they used Sicilian more within and outside their homes than bilectal HS in Germany.

3.3 Procedure

The data collection was carried out online using SoSci Survey (Leiner 2014). Speakers were recruited via personal contacts, the distribution of flyers in universities and social media. The procedure started with an introduction to the components of the survey. Once participants had given their consent, they were asked for biographical data, then carried out the Sicilian vocabulary task (LexSIC), filled in some information on dialect use and attitudes and finally completed the Italian vocabulary task (DIALANG). The entire questionnaire took about 40 minutes to complete.

The LexSIC began by explaining to the task-taker that they would read and hear verbs which were either Sicilian or invented, and that their task was to indicate whether they *knew the meaning* of the word (yes) or not (no) (see Appendix 2). The same instruction was used for the DIALANG. In our instructions we deviated from the original instructions of the DIALANG. In the DIALANG, the test taker is asked to indicate whether they thought a verb existed or not. Due to previous experiences in an online study where this instruction was felt to invite guessing behavior, and since regional variation was an additional factor here (i.e., participants being used to variation and being more tolerant), we opted for the instruction to indicate one's knowledge of the verb. Moreover, in order to prevent participants from being overly strict and reject everything that does not exist within their own local variety, we reminded them of regional variation and instructed them to accept the word if they thought that it was a regional variant of a word they knew. The task began with three practice items, two real words and one pseudo-word.

4. Towards LexSIC

4.1 Item assessment

Item assessment was done on the basis of both CTT and IRT. One main difference between the approaches is that item analysis in CTT is sample-dependent (Bachman 2004), where the data underlying the test score analyses come from a specific group of individuals, and generalizing to an underlying population is technically not possible.

[5] The German "Middle School" degree (*Realschule*) is an intermediate degree between *Hauptschule* (lowest degree, 9 years of school) and *Gymnasium* (highest degree, 13 years of school). It corresponds to the Italian degree after *scuola media*. In our analysis, we combined Haupt- and Realschule due to the low number of participants in these categories.

In contrast, IRT is a probability-based framework, and estimates of item and test-taker characteristics are independent of the particular sample at hand. Another difference is the possibility afforded by IRT to have item difficulty estimates and person[6] ability estimates projected on the same scale. In CCT, it is only possible to draw on the average performance of a specific group for calculations of the difficulty index for items.

The original test consisted of 103 items (65 real words and 38 pseudo-words). Following the procedure of previously created tests with a similar goal, we carried out separate analyses for real words and pseudo-words (e.g., Lemhöfer & Broersma 2012; Brysbaert 2013; Amenta et al. 2020; Salmela 2021).

As a first step, we excluded participants from the dataset who had a distinctly high number (above 50%) of false alarms (i.e., the identification of a pseudo-word as a real word) in either the LexSIC or the Italian vocabulary task taken from the DIALANG test battery. Of the original sample of 100 participants, 10 participants were excluded due to guessing behavior in the LexSIC (n = 7) or in the DIALANG (n = 3). This led to the removal of 7 participants (5 Sicilians and 2 HS) for assumed guessing behavior in the LexSIC and 3 participants (2 Sicilians and 1 HS) for assumed guessing behavior in the DIALANG component.

We next ran an item-total correlation analysis[7] on the real words and removed the 9 items with the lowest coefficient value (anything under .3). This led to the exclusion of the following items: *crisciri* (it. *crescere* 'grow'), *curriri* (it. *correre* 'run'), *dormiri* (it. *dormire* 'sleep'), *furriari* (it. *girare, andare in giro* 'go around'), *leggiri* (it. leggere 'read'), *meravigghiari* (it. *meravigliare* 'wonder'), *rrinuvari* (it. *rinnovare* 'renew'), *rumpiri* (it. *rompere* 'break'), *vinciri* (it. *vincere* 'win'). All these items closely resembled the corresponding Italian words. An exception is *furriari*, which was removed because there was an item with the same meaning in the task.[8] We then looked at the data fit using Rasch Analysis with the goal of eliminating 6 additional items to arrive at a total of 50 real words. In selecting items to be removed, we followed a combined strategy of (i) eliminating items with the same Rasch estimated item difficulty level, (ii) considering fit to the model (infit mean square values), and (iii) considering how close these words were to their equivalents in Italian. This led to the exclusion of *amari* (it. *amare* 'love'), *cuntari* (it *contare* 'count'), *lavari* (it. *lavare* 'wash') and *sputari* (it. *sputare* 'spit'). We further excluded *spiari* (it. *chiedere* 'ask'), and *viviri* (it. *bere* 'drink') because they can be considered "false

[6] Following IRT jargon, we use the terms "person ability" and "person values", where "person" is equivalent to "participant".

[7] This is the correlation of each item with the total score for this item. If everything works, the items should discriminate between those participants who have a high score and those who have a low score.

[8] This inclusion of the two items *furriari* and *firriari* was preceded by a long discussion amongst speakers of Sicilian, who were of the opinion that these two items were dialectally in complementary distribution, and that one might be perceived as more archaic than the other (making it "easier" for HS who had been living in the diaspora for more than 50 years). Since this discussion was based on subjective intuition, we nevertheless decided to keep both in the dataset for the piloting to see how participants would treat them. Contrary to expectations, the judgments for these two items were not complementary, as some speakers accepted both and there was clear acceptance bias in Sicilians vs. HS. Since both ended up amongst the items with the best item fit, we decided to keep one, namely the one with the better fit.

friends" of it. *spiare* 'spy' and it. *vivere* 'live'. Although in the latter two cases, the translation equivalents (*viviri-bere, spiari-chiedere*) look very different from each other, participants with a lower proficiency might (erroneously) indicate that they know the meaning due to the existence of quasi-homophones in Italian.

For the 38 pseudo-verbs we also ran item total correlation, and we removed the items with the lowest coefficient value (items under .3). This led to the removal of 3 items (*prelari, stifiddari, talamari*). In order to arrive at 25 pseudo-words, we then used a combined strategy of (i) eliminating items with the same Rasch estimated item difficulty level, while (ii) considering fit to the model (infit mean square values). This led to the elimination of *beddari, biggiari, crundari, farrari, undari, positari, pratiari, piddari, libbari, taddicari*. The final set of 75 items was used to calculate a score (the LexSIC score).

4.2 Task assessment and predictions

To validate the test, we compared the results across groups, and we ran correlations between the LexSIC scores with self-reported proficiency, dialect use, age, education and the DIALANG scores in Italian. We followed the same scoring procedure for the Italian DIALANG placement test and for the LexSIC in order to ensure comparability across varieties. This means that for the DIALANG, rather than following the procedures of the online instructions, we used the original items but implemented them in our own online test battery.

We expected participants in Sicily to have the highest scores and participants from North Italy to have the lowest. We further expected positive correlations between the LexSIC score and self-assessed proficiency as well as dialect use. Since dialect use tends to be less frequent with higher education levels and younger age, we expected lower LexSIC scores with increasing educational levels. Finally, we wanted to see how the LexSIC scores relate to DIALANG scores because dialect use is often purported to have negative consequences for the standard language. An alternative possibility is that there are general benefits for acquiring two closely related dialects, just as there are general benefits of bilingual experience (Bialystok et al. 2004). Minimally, it stands to reason that the additional acquisition of closely related languages does not compromise language acquisition in the majority language, as has been shown by Garraffa et al. (2015, 2017) for Sardinian-Italian bilinguals.

5. Results

An IRT Rasch analysis of the final 50 real word items selected through the previous analyses resulted in the values shown in Table 2 below. A Wright map is provided in Appendix A3. The summarized person and item values in Table 2 indicate that the item and person infit and outfit statistics were within the expected range (MNSQ 0.5-1.5, Z ±2) (Green 2013). The person separation index reveals how well the items separate persons measured, whereas the item separation index shows how well a tested sample of people is able to separate the items. Linacre (2002) has suggested that item separation indices of 3 or greater are desirable. For the person separation index, 1.5 is conventionally acceptable, 2.0 is good, and 3.0 is excellent. In our case the item separation of 3.01 matches the desired threshold of 3, and the person separation of 2.22

is good. Furthermore, a Cronbach's alpha coefficient was observed at .97, which together with the item reliability of .90 suggests that the replicability of the item performance on a similar test population is sufficiently high. The mean item-total correlation was .60 (all items but one reached > .30, range .23-.84). Thus, the items included in the final version of LexSIC have a good discriminative power.

Table 2. Person and item measure statistics

	Measure		**Infit**		**Outfit**		**Separation**	**Reliability**
	Mean	***SE***	***MNSQ***	***ZSTD***	***MNSQ***	***ZSTD***		
Person	3.35	.75	.95	.04	.93	.04	2.22	.83
Item	.00	.48	.97	-.05	1.01	-.14	3.01	.90

The Wright map in Appendix A3 indicates that the final version set of items matches the abilities well for the lower segment, but that there are few items that can match the persons with the very highest abilities (Logits of > 5).

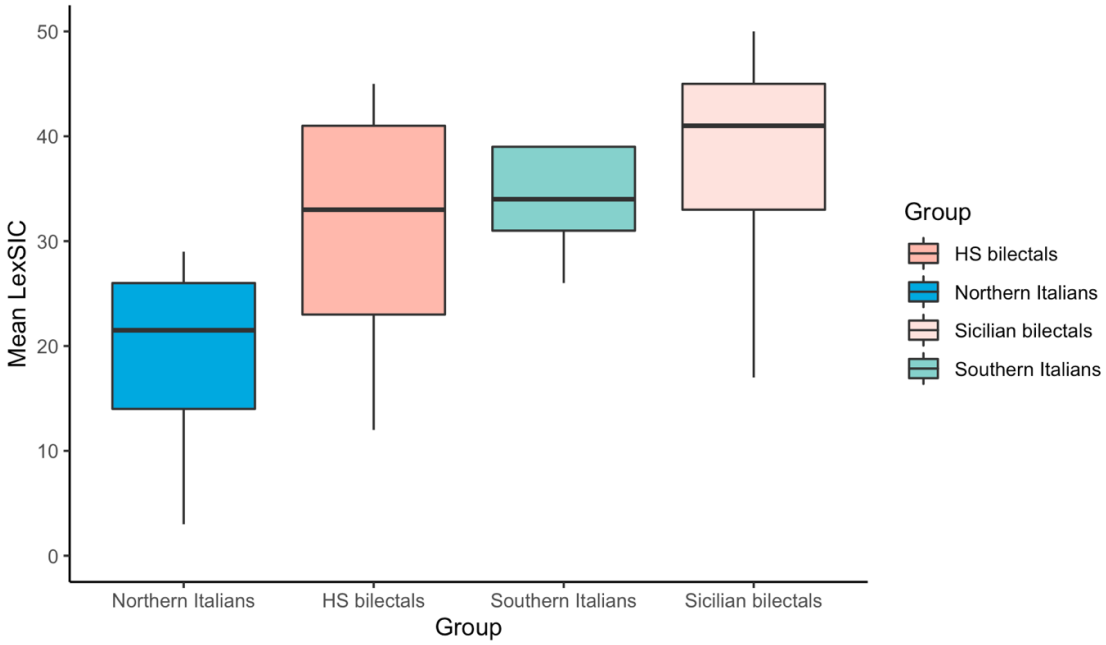
We next wanted to assess how LexSIC scores can differentiate between the different groups. Following Brysbart (2013) and Amenta et al. (2020) we calculated the test score as in (1). This formula penalized guessing behavior (Izura et al. 2014). There is a maximum score of 50, which can only be obtained by saying "yes" to all the words and to none of the non-words.

(1) LexSIC Score = $N_{\text{yes to words}} - 2 * N_{\text{yes to nonwords}}$

The results on the LexSIC for the four groups of participants are shown in Figure 1. As expected, Northern Italians had the lowest scores (M = 19.50, SD = 8.14, *range* = 3-29) and speakers from Sicily scored highest (M = 38.41, SD = 7.55, *range* = 17-50). Speakers of other South Italian dialects scored within the range of Sicilians, though somewhat lower on average (M = 33.80, SD = 4.96, *range* = 26-39). The HS bilectals showed the widest range of scores (M = 31.47, SD = 10.17, *range* = 12-45), together with bilectal Sicilians.

In order to assess whether the differences between groups were significant, we ran a poisson generalized linear mixed effect model (*glm* function from the *stats* package in R (R Core Team 2022) with LexSIC score as dependent variable and Group as fixed effect (Model 1). We found a significant effect of Group (X^2 = 47.5, p < .001) and the post-hoc test with Bonferroni correction (*emmeans* function, Length 2023) confirmed significant differences between the North Italians and the HS bilectals (β = 11.97, SE = 3.25, p = .002), as well as between the North Italians and the Sicilian bilectals (β = -18.91, SE = 2.85, p < .0001) and the other Southern Italians (β = - 14.3, SE = 4.55, p = .014). The Southern Italian group did not differ significantly from neither the HS bilectal group (β = - 2.33, SE = 4.18, p = 1) nor the Sicilian bilectal group (β = 4.61, SE = 3.88, p = 1). The difference between the HS bilectals and the Sicilian bilectals was significant (β = - 6.94, SE = 2.21, p = .014).

Figure 1. LexSIC score by group. The maximum LexSIC score was 50



Following Salmela et al. (2021), we next ran a set of additional analyses to explore how the performance in LexSIC was related to some additional measures, i.e., receptive skills, use of dialect within the family, age as well as the participants and their parents' education level; we also included their DIALANG (i.e., Italian proficiency) score. These additional analyses were run for a subset of the participants, i.e., only for the Sicilian and the HS bilectals (n = 75).

First, we generated a correlation matrix (Table 3) to examine the relationship between the additional predictors and the participants' performance in our vocabulary task. The data on receptive proficiency was provided by the speakers themselves on a scale from 1-6. Since many of our participants reported only receptive and no productive skills, we based our comparison on the former. For use within the family, we calculated a composite score. The Sicilian bilectals were asked to indicate on a 5-point scale to what extent they used Sicilian and German, respectively (0 = only Italian, 4 = only Sicilian). This question was asked for different relationships, including parents, grandparents, siblings, cousins, aunts and uncles, and partner. The scores were added up. If speakers chose "does not apply", the answer was coded as "0". The HS in Germany answered the same questions, comparing their use of Sicilian vs. Italian. However, they were asked for an additional comparison of how much German they used compared to the two Romance varieties (0 = only German, 4 = only Italian/Sicilian). The Sicilian use score for this group was then calculated as a portion of the Sicilian/Italian (vs. German) score. This accounts for the different situation of minority speakers abroad who have fewer opportunities of using their minority languages.[9] We did not take situations of language use outside of the home into

[9] We are aware that asking for *relative use of Italian vs. Sicilian on one scale* is not precise because language use is not complementary such that every time someone uses Italian, they do not use Sicilian. Instead, there are contexts in which several languages can be used

account because all groups indicated few such opportunities. The participants' education, their mothers' and fathers' education were scored on a scale from 1 to 7.

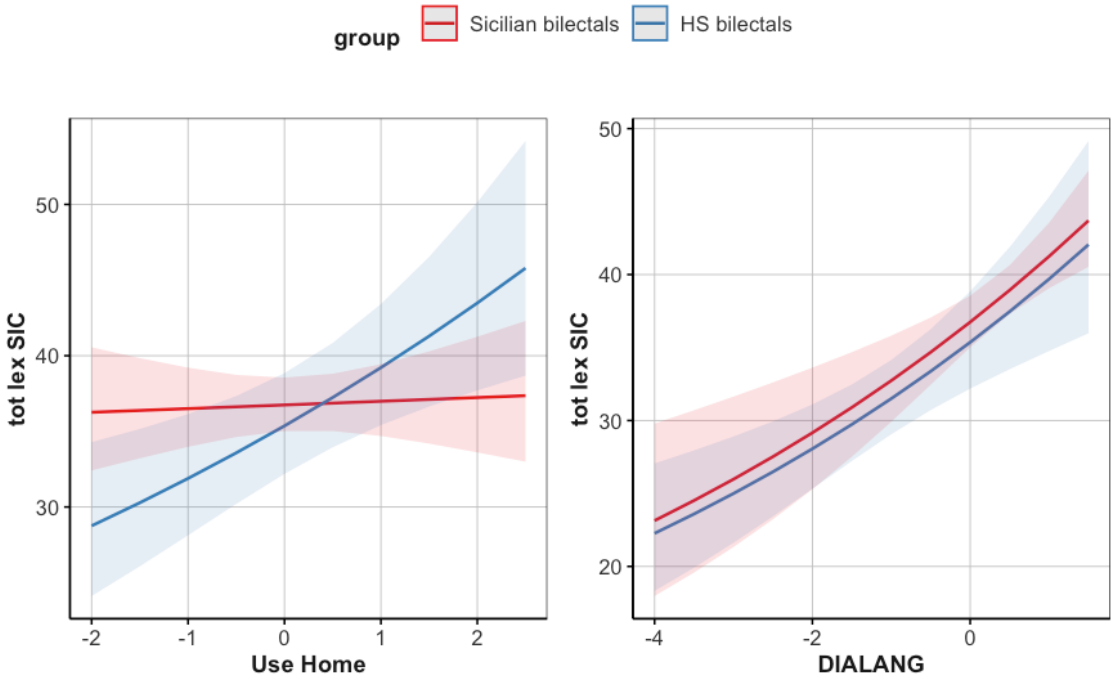
Table 3. Correlation matrix displaying rho and p-values for pairwise Spearman's correlations. Statistical significance is indicated as follows: r* ($p < 0.05$), r** ($p < 0.01$), and r*** ($p < 0.001$)

	LexSIC	Age	Education	Edu. (mother)	Edu. (father)	Receptive skills	Use
Age	.27*						
Education	.16	.18					
Edu. (mother)	- .03	- .26*	.38***				
Edu. (father)	.12	- .15	.37***	.82***			
Receptive skills	.24*	.25*	- .06	- .06	- .11		
UseHome	.31**	.32**	- .30**	- .36**	- .35**	.52***	
DIALANG	.50***	.43***	.28*	.13	.25*	.20	.10

To examine the predictive ability of the independent variables reported in Table 3 on the participants' performance in the LexSIC, we built a generalized linear model (Model 2) with stepwise regression using a starting model that predicted LexSIC scores as a function of UseHome, Education, DIALANG scores, and participants' Age. All predictors were scaled, and they were included in the model in a two-way interaction with Group. Given the high correlation between use within the home and reported receptive language skills ($r = 0.52$, $p < .001$) (see Table 3), as well as between participants' education and their mothers' ($r = 0.38$, $p < .001$) and fathers' ($r = 0.37$, $p < .001$) education, receptive skills and parental education were not included in the model. We used the *drop1* function in R (*stats* package, R Core Team 2022) to remove interactions and/or predictors that did not significantly improve the model fit. Our best-fit model included DIALANG scores and a two-way interaction between Use at home and Group. All other predictors and interaction were not significant. The results indicated a significant main effect of Use ($X^2 = 9.6$, $p = 0.001$) and DIALANG scores ($X^2 = 16.43$, $p < 0.001$), as well as a significant interaction between Group and UseHome ($X^2 = 5.31$, $p = 0.02$). To test how the interaction plays out, a post-hoc analysis was performed using the *emtrends* function from the *emmeans* package in R (Lenth 2023). The analysis revealed that Sicilian bilectals and HS bilectals differ significantly in the context of their interaction with UseHome ($\beta = 0.09$, $SE = 0.04$, $p = 0.02$). The results of the model are shown in Figure 2. The left panel illustrates how the use of dialect within the family affects LexSIC performance: HS bilectals show a significant improvement in vocabulary task performance when exposed to increased dialect use at home ($\beta = 0.1$, $SE = 0.03$, $p = 0.002$), while this is not the case for Sicilian bilectals ($\beta = 0$, $SE = 0.02$, $p = 0.7$). On the right panel, the effect of DIALANG scores on LexSIC scores is illustrated, showing for both groups that higher proficiency in Italian significantly predicts higher proficiency in Sicilian. Detailed results for Model 1 are presented in Tables 4 (model output) and Table 5 (post-hoc analysis), while the model output and the post-hoc analysis results of Model 2 are presented in Table 6 and 7 respectively, which can be found in the Appendix (A4).

and contexts where languages beyond Italian and Sicilian are used. Nevertheless, we chose this method because it has been used in dialect studies, and it is perceived to be easier for participants to compare relative use in two varieties rather than, e.g., estimating the number of hours per week that they use each variety.

Figure 2. Model 2 output. Effect of **Use** at home (left) and **DIALANG** scores (right) on LexSIC.



6. Discussion

6.1 LexSIC as a placement test for bilectal proficiency

We developed and provided initial validation of LexSIC, an online vocabulary assessment tool for Sicilian. The test was modeled after the DIALANG placement test with adaptations in the instructions and the presentation of stimuli, which were presented in both oral and written mode. Instead of comparing L1 and L2 speakers, as, e.g., in Lemhöfer and Broersma (2012), we compared bilectals in the home country (Sicily) with bilectal HS in the diaspora (Germany) and with speakers of related and unrelated varieties. The LexSIC scores were highest amongst the Sicilian (homeland) bilectals and speakers of related varieties. They were lower and highly variable in the HS bilectals and lowest in speakers from North Italy, who were expected to have the least contact with Sicilian.

To initially validate the LexSIC data, we ran correlations with other factors that were shown to relate to vocabulary scores in previous studies. These correlations focused on the two groups who had actual experience with Sicilian because they were ethnic Sicilians, i.e., the bilectal Sicilians and the heritage bilectals from Germany. Like in previous L2 studies (e.g., Amenta et al. 2021; Salmela et al. 2021), the LexSIC scores correlated strongly with self-assessed proficiency, although in our study we focused on receptive proficiency because we included participants without productive Sicilian skills. The LexSIC scores further correlated with reported use of Sicilian within the home. These results support the validity of LexSIC, and they are generally in line with the finding that language experience influences vocabulary size. However, a more comprehensive statistical analysis revealed that the correlation between

LexSIC and use was restricted to the HS bilectals in Germany (we could not test proficiency in the same model, as the two variables were highly correlated). Why did we not see such an effect in the Sicilian bilectals? One potential explanation is that in Sicily, even for those who report to have less proficiency and use less Sicilian at home, the communal use might be sufficient to acquire the meaning of a large number of our verb items. By contrast, HS bilectals have no passive immersion in the society. Their opportunities to gain language experience are restricted to the home context, and this is why the amount of exposure to Sicilian in the family can more easily change the results, i.e., have a significant impact on the development of dialect proficiency.

As to age and education, we expected older participants with the lowest educational level to know the dialect best, because this is where dialect use has been consistently reported to be most widespread (Grassi et al. 1997; Berruto 2018). However, this prediction was not confirmed. Age showed a weak correlation with LexSIC scores (Table 3), but when included in the model with other predictors the effect disappeared, which seems to run counter to the expectation that dialect use is more widespread amongst the older generation. Similarly, there was no correlation of LexSIC scores with education. This unexpected absence of educational effects might be explained by the relatively high number of academics with linguistic training in our sample. Another possible explanation, pointed out to us by a reviewer, is that presently, many university degree holders are also first-generation degree holders, potentially coming from backgrounds where Sicilian was spoken. Alternatively, the prevalence of dialect skills reported in previous studies for speakers with lower education might be in the process of changing, as a result of changing attitudes towards a more positive perception of dialectal competence. We are thus inclined to interpret the absence of relations between the LexSIC scores and education as a combined effect of the (possibly somewhat biased) sample together with positive effects of a reevaluation of dialects, as discussed in Parry (2010) and Berruto (2018). As Parry (2010) has pointed out, while the ability to speak Italian used to be an indicator of higher education, this marker has become less meaningful today because everyone speaks Italian confidently. As a consequence, regardless of educational level, speaking the dialect has become more acceptable. At the same time, we cannot exclude that our sample is over-representative of people with a positive attitude towards Sicilian, because a survey of this type is less likely to generate interest and readiness to participate amongst people with lower Sicilian skills and negative attitudes.

Finally, it is interesting to see that higher DIALANG scores (a.k.a. better Italian skills) correspond with higher LexSIC skills. This result contributes to dismantling the myth that speaking a regional language/ dialect is detrimental for the development of the standard language, as has been shown very nicely by Garraffa et al. (2015, 2017) for Sardinian-Italian bilectals. It has become obvious that human brains can easily cope with multilingualism and, moreover, that there is some kind of cumulative enhancement when learning more than one language.

6.2 Methodological challenges for vocabulary tests in dialects/ regional minority languages

In the introduction we have pointed to challenges pertaining to the creation of proficiency assessment instruments in regional minority languages/ dialects, including variation and lack of standardization. An additional concern resulting from the

validation of LexSIC is that we ended up with a task that is felt to be "too easy". One indication appears to be that the person-item map for real verbs in the Appendix A3 shows few items and many participants in the upper range. Another potential indication is that northern Italians, whose language/s are very distinct from Sicilian, have scored around 20. Despite removing many of the "easier" items during the item assessment stage, the final test version, as the list of items in the Appendix A1 shows, still contains some items that may be easily recognizable by speakers of Italian who have never been to Sicily, e.g., *aviri* (it. *avere* "have"), *nivicari* (it. *nevicare* 'snow'), *vuliri* (it. *volere* 'want'). This raises question is whether it is due to typological proximity, viz. the cognate status of most items, that the participants can easily identify Sicilian words, and whether this compromises the validity of LexSIC as a suitable test of Sicilian proficiency. For the following reasons, we do not believe that is a major concern for our task: First, although North Italians do not score at 0, their scores are significantly different from those of Sicilians, thus underlining that our test is successful in discriminating between people who are exposed to Sicilian and people who are not. Second, it is natural that Italians can understand a good amount of speech when listening to Sicilian, just as Germans can understand something if they go to the Netherlands or Italians when going to Spain, simply because the languages are typologically very close. Thus, the removal of all items whose meaning can be guessed due to knowledge of Italian would compromise ecological validity. Third, many of the items in the task denote activities that are part of daily life. While this means that that technical terms or high register are underrepresented, this is where Sicilian is mostly used. Technical terms are typically borrowed from other languages (Italian or English), and the high variety is Italian rather than Sicilian. Thus, excluding technical terms and high register contributes to the ecological validity of the test. Finally, a concern related to typological vicinity is that this is what explains the correlation between LexSIC and DIALANG scores, rather than beneficial effects of bilingualism/bilectalism (Figure 5). We cannot exclude this possibility, but we consider it unlikely, because the items in the two tests are not translation equivalents.[10]

There are three additional methodological points we wish to discuss: the number of test takers (n = 100), the number of final items (n = 75) and the choice of verb items (rather than several word classes). As to the first point, it is true that, ideally, validation could have included a larger number of control speakers to support the finding that speaking Italian does not automatically result in receptive knowledge of any regional variety spoken in Italy. Such evidence could also come from similar language pairs; a follow-up study creating a test for Venetan with a larger number of control speakers from Lombardy and Sicily is in the work. Second, we chose to end up with a final list of 50 real words and 25 pseudo-words to ensure comparability with the DIALANG placement tests. There is no doubt that this number is low for a proper vocabulary size test (Gyllstad et al. 2021). However, we wish to remind the reader that our goal was not to create a vocabulary size test but a placement test. That is, we are not primarily interested in an estimate of the test taker's vocabulary size but in an estimate of where on the proficiency scale they are compared to other test takers and compared to their own proficiency in Italian. Moreover, we aimed for a test that is as

[10] Another potential explanation for this correlation is that some participants are simply better at taking tests. We cannot exclude this possibility either, but the observed correlation between language use scores and LexSIC scores speaks against it.

quick as possible, because, unlike in many L2 studies, the typical target group of LexSIC will take a placement test in at least one additional variety and possibly other tests assessing other properties.

The third point for discussion concerns restricting the task to verb items. In doing so, we followed the DIALANG design and deviated from LexITA, which has also included nouns and adjectives. The advantage of verbs is that they are typically later acquired (at least than nouns) and are arguably more central to the clause as they determine the number and semantic role of their arguments. Moreover, nouns are more likely than verbs to display diatopic variation within regional languages, which would have been a disadvantage for our task, as it is supposed to work for all Sicilian speakers regardless of their geographical origin. This said, we do not think that using adjectives and nouns is generally disadvantageous, especially if used in combination with other lexical class members. On the contrary, it might have given us a richer palate to choose items from since were already constrained in terms of register. As mentioned above, the major reason for selecting verbs only was to ensure comparability with placement tests that are already available in different languages.

A remaining challenge in assessing proficiency in regional minority languages is that there may always be some variation in outcomes even at higher proficiency levels since regional languages vary more than standard languages. This means that, although we have tried to avoid the inclusion of items that are area-specific to the extent possible, it is likely for a test-taker to encounter items that do not exist in their variety. Moreover, speakers of a minority language may be more tolerant towards variation and thus more tempted to accept pseudo-verbs. While we are aware of this potential caveat, we do not think that the consequence should be to refrain from assessing proficiency in regional languages. Instead, potential users of such assessment tools should be cautious when interpreting their data and look into factors that might explain the potentially found variation.

Finally, based on data from a few speakers outside of Sicily, we suspect that the test is not ideal to assess the proficiency of speakers in Calabria and Puglia, even though their varieties are closely related to those spoken in Sicily and part of the same continuum. Although these speakers could correctly identify the majority of the items, it seems unfair to test them on a task that was tailored to the varieties in Sicily. However, we have had only five such participants in our sample, and anyone interested is invited to use the existing items, replace the ones that would be problematic and create a new version of the task.

7. Availability and Conclusions

LexSIC is a reliable and quick assessment of Sicilian vocabulary that arose from the need to assess language proficiency in heritage speakers beyond the standard language. The completion of the test requires about five minutes, and it can be used as a pen and paper test or on a computer/smartphone/tablet. We administered our items in a randomized order and we have presented them in a bimodal fashion (aurally and written). If randomization is impossible for technical reasons, we suggest creating two or more semi-randomized versions. The final list of the LexSIC items and the instruction we have used are included in the Appendix. For everyone who would like

to use a bimodal version, the recordings of the individual items are available through Iris.[11]

A particularly interesting finding during the initial validation was that the LexSIC scores correlated more strongly with other language use measures in the group of HS bilectals than in the group of Sicilian bilectals. This might point to an important difference between heritage bilingualism in the diaspora and autochthonous (indigenous) heritage bilectals: Language development in the diaspora setting is very much dependent on input at home, and every single opportunity to gain experience with Sicilian makes a significant difference. By contrast, speakers in the homeland (Sicily) might have sufficient experience with Sicilian outside of their homes to develop good receptive proficiency, even if within the families no extra efforts are made to promote Sicilian. Of course, by making this observation, we do not mean to imply that there is no necessity for homeland bilectals to use their minority language/ dialect at home, because we have only been concerned with receptive proficiency, and it is conceivable that the development of productive proficiency requires more. Second, Sicilian is considered to be stable; the situation may play out differently in cases where a language or dialect is at a more advanced stage towards endangerment.

Acknowledgments

We wish to thank our colleagues, friends and family who supported the data collection. This research did not receive any specific grant from funding agencies.

References

Alderson, J. Charles. 2005. *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.

Alderson, J. Charles & Ari Huhta. 2005. The development of a suite of computer-based diagnosis tests based on the Common European Framework. *Language Testing* 22(3): 301–320. https://doi.org/10.1191/0265532205lt310oa

Amenta, Simona, Badan, Linda, & Marc Brysbaert. 2021. LexITA: A Quick and Reliable Assessment Tool for Italian L2 Receptive Vocabulary Size. *Applied Linguistics* 42(2): 292–314. https://doi.org/10.1093/applin/amaa020

Anderson, Richard C. & Peter Freebody. 1983. Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutton (ed.), *Advances in Reading/Language Research*, Vol. 2, 231–256. Greenwich, Connecticut: JAI Press.

Andriani, Luigi, Casalicchio, Jan, Ciconte, Francesco M., D'Alessandro, Roberta, Frasson, Alberto, Van Osch, Brechje, Sorgini, Luana, & Silvia Terenghi. 2022. Documenting Italo-Romance minority languages in the Americas. Problems and tentative solutions. In: A. Nevins & M. Cole (eds.), *Contemporary research in minority and diaspora languages of Europe*, 9–56. Berlin: Language Science Press.

[11] https://www.iris-database.org/details/fzcgn-9e1R5

Bachman, Lyle. 2004. *Statistical analyses for language assessment*. Cambridge University Press.

Bates, Elizabeth & Judith C. Goodman. 1999. On the emergence of grammar from the lexicon. In B. Macwhinney (ed.), *The emergence of language*, 29– 79. Mahwah, NJ: Erlbaum.

Beglar, David. 2010. A Rasch-based validation of the Vocabulary Size Test. *Language Testing* 27(1): 101–118. https://doi.org/10.1177/0265532209340194

Benmamoun, Elabbas, Montrul, Silvina, & Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39(3-4): 129-181. https://doi.org/10.1515/tl-2013-0009

Berruto, Gaetano. 1987. *Sociolinguistica dell'italiano contemporaneo*. Roma: La Nuova Italia Scientifica. New updated edition in 2012.

Berruto, Gaetano. 1989. Main topics and findings in Italian sociolinguistics. *International Journal of the Sociology of Language* 76: 7–30. https://doi.org/10.1515/ijsl.1989.76.7

Berruto, Gaetano. 2018. The languages and dialects of Italy. In W. Ayres-Bennett & J. Carruthers (eds.), *Manual of Romance Sociolinguistics*, 494–525. Boston, Berlin: De Gruyter. https://doi.org/10.1515/9783110365955-019

Besler, Alexandra, Ferin, Maria, Venagli, Ilaria, & Tanja Kupisch. in prep. On the relation between attitudes and dialect maintenance (Sicilian and Venetan) in Italy, Ms. University of Konstanz.

Bialystok, Ellen, Craik, Fergus I. M., Klein, Raymond, & Mythili Viswanathan. 2004. Bilingualism, aging, and cognitive control: evidence from the Simon task. Psychology and Aging 19: 290–303. https://doi.org/10.1037/0882-7974.19.2.290

Brysbaert, Marc. 2013. Lextale_FR A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica*, 53(1): 23–37. https://doi.org/10.5334/pb-53-1-23

Coluccia, Rosario. 1994. Il volgare nel Mezzogiorno. In L. Serianni & P. Trifone (eds), *Storia della lingua italiana. III: Le altre lingue*, 373-405. Turin: Einaudi.

Council of Europe. 1992. The European Charter for Regional or Minority Languages. www.coe.int/en/web/european-charter- regional-or-minority-languages/text-of-the-charter.

Cruschina, Silvio. 2020. The classification of Sicilian dialects: Language change and contact. *L'Italia Dialettale* 81: 79–103.

De Mauro, Tullio. 2005. Storia linguistica dell'Italia unita, 9. ed. Roma: GLF Editori Laterza.

De Renzo, Francesco. 2008. Per un'analisi della situazione sociolinguistica dell'Italia contemporanea. Italiano, dialetti e altre lingue. Italica 85(1): 44-62.

Eberhard, David M., Simons, Gary F., & Charles D. Fennig (eds.). 2022. *Ethnologue: Languages of the World.* Twenty-fifth edition. SIL International. http://www.ethnologue.com

Ferguson, Charles A. 1964. Diglossia. In D. Hymes (ed.), Language in Culture and Society, 425–439. New York: Harper & Row.

Garraffa, Maria, Beveridge, Madeleine, & Antonella Sorace. 2015. Linguistic and cognitive skills in Sardinian-Italian bilingual children. Frontiers in Psychology 6(1898): 1–15. https://doi.org/10.3389/fpsyg.2015.01898

Garraffa, Maria, Obregon̦, Mateo, & Antonella Sorace. 2017. Linguistic and cognitive effects of bilingualism with regional minority languages: a study of Sardinian-Italian adult speakers. Frontiers in Psychology 8(1907): 1-11. https://doi.org/10.3389/fpsyg.2017.01907

Gathercole, Susan E. & Alan D. Baddeley. 1989. Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language* 28: 200–213. https://doi.org/10.1016/0749-596X(89)90044-2

Grassi, Corrado, Sobrero, Alberto A., & Tullio Telmon. 1997. *Fondamenti di dialettologia italiana*. Rome: Laterza.

Green, Rita. 2013. *Statistical analyses for language testers*. London: Palgrave Macmillan.

Gyllstad, Henrik. 2013. Looking at L2 vocabulary knowledge dimensions from an assessment perspective – Challenges and potential solutions. In C. Bardel, B. Laufer, & C. Lindqvist (eds.), *L2 Vocabulary Acquisition, Knowledge and Use: New Perspectives on Assessment and Corpus Analysis,* 11-28. Eurosla Monographs Series, 2. EUROSLA.

Gyllstad, Henrik, Vilkaitė, Laura, & Norbert Schmitt. 2015. Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. ITL *International Journal of Applied Linguistics* 166(2): 278–306. https://doi.org/10.1075/itl.166.2.04gyl

Gyllstad, Henrik, McLean, Stuart, & Jeffrey Stewart. 2021. Using confidence intervals to determine adequate item sample sizes for vocabulary tests: An essential but overlooked practice. *Language Testing* 38(4): 558–579. https://doi.org/10.1177/0265532220979562

Haller, Hermann W. 1993. *Una lingua perduta e ritrovata. L'italiano degli italo-americani*. Firenze: La Nuova Italia.

Haller, Hermann W. 1997. The dialects abroad. In M. Maiden und M. Parry (eds.), *The Dialects of Italy*, 401–411. London: Routledge.

Haller, Hermann W. 2014. Evolving linguistic identities among the Italian-American youth: Perceptions from linguistic autobiographies. *Forum Italicum* 48(2): 238–252. https://doi.org/10.1177/0014585814529230

Hilton, Heather. 2008. The link between vocabulary knowledge and spoken L2 fluency. *The Language Learning Journal 36*: 153–166. https://doi.org/10.1080/09571730802389983

ISTAT. 2007. *La lingua italiana, i dialetti e le lingue straniere. Anno 2006,* online, portal-lem.com/images/it/Italie/Lingue_e_dialetti_e_lingue_straniere_in_Italia.pdf.

ISTAT. 2012. *L'uso della lingua italiana, dei dialetti e di altre lingue in Italia*, published online 27th October 2014, http://www.istat.it/archivio/136496.

Izura, Christina, Cuetos, Fernando, & Marc Brysbaert. 2014. Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica* 35(1): 49–66.

Jeon, Eun H. & Junko Yamashita. 2014. L2 reading comprehension and its correlates: A Meta-analysis. *Language Learning* 64: 160–212. https://doi.org/10.1111/lang.1203

Laufer, Batia & Geke C. Ravenhorst-Kalovski. 2010. Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1): 15–30.

Leiner, Dominik J. 2014. SoSci Survey (Version 2.4.00-i). [Computer software]
Lenth R (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.4-1, <https://CRAN.R-project.org/package=emmeans>.

Leivada, Evelina, Papadopoulou, Elena, Kambanaros, Maria & Kleanthes Grohman. 2017. The Influence of Bilectalism and Non-standardization on the Perception of Native Grammatical Variants. *Frontiers in Psychology* 8: 205.

Lemhöfer, Kristin & Mirjam Broersma. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44: 325–43. https://doi.org/10.3758/s13428-011-0146-0

Linacre, John Michael. 2002. What do infit and outfit, mean-square and standardized mean. *Rasch Measurement Transactions* 16(2): 878.

Loporcaro, Michele. 2009. *Profilo linguistico dei dialetti italiani*. Bari: Editori Laterza.

Matranga, Vito & Roberto Sottile. 2013. La variazione dialettale nello spazio geografico. In G. Ruffino (ed.), *Lingue e culture in Sicilia*, 215–274. Palermo: Centro di studi filologici e linguistici siciliani.

McLean, Stuart, Kramer, Brandon, & Jeffrey Stewart. 2015. An empirical examination of the effect of guessing on vocabulary size test scores. *Vocabulary Learning and Instruction* 4(1): 26–35.

Meara, Paul. 1996. The dimensions of lexical competence. In G. Brown, K. Malmkjær, & J. Williams (eds.), *Performance and competence in second language acquisition*, 33–55. Cambridge: Cambridge University Press.

Meara, Paul & Barbara Buxton. 1987. An alternative to multiple choice vocabulary tests. *Language Testing* 4: 142-154.

Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.39

Montrul, Silvina. 2016. *The Acquisition of Heritage Languages*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139030502

Moseley, Christopher. 2010. *Atlas of the worlds' languages in danger*. Unesco Publishing.

Nation, Paul. 1993. Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (eds.), *The bilingual Lexicon*, 115–134. Amsterdam: John Benjamins.

Parry, Mair. 2010. Matters of choice: Language preferences in Italy today. *The Italianist* 30(2): 61-77.

Pellegrini, Giovan Battista. 1977. Carta dei dialetti d'Italia (Profilo dei dialetti italiani 0). Pisa: Pacini**.**

Polinsky, Maria. 2018. *Heritage Languages and their Speakers*. Cambridge: Cambridge University Press.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rubino, Antonia. 2014. *Trilingual Talk in Sicilian-Australian Migrant Families: Playing Out Identities Through Language Alternation*. London: Palgrave Macmillan.

Ruffino, Giovanni. 1991. *Dialetto e dialetti di Sicilia.* Palermo: CUSL.

Ruffino, Giovanni. 2001. *Profili linguistici delle regioni. Sicilia*. Bari: Laterza.

Ruffino, Giovanni. 2006. *L'indialetto ha la faccia scura. Giudizi e pregiudizi linguistici dei bambini italiani.* Palermo: Sellerio.

Salmela, Rosa, Lehtonen, Minna, Garusi, Stefano, & Raymond Bertram. 2021. Lexize: A test to quickly assess vocabulary knowledge in Finnish. *Scandinavian Journal of Psychology* 62(6): 775–894. https://doi.org/10.1111/sjop.12768

Schmid, Marc. 2014. *Italienische Migration nach Deutschland. Soziohistorischer Hintergrund und Situation im Bildungssystem.* Wiesbaden: Springer VS.

Schmitt, Norbert. 2014. Size and depth of vocabulary knowledge: What the research shows. *Language Learning* 64: 913–951. https://doi.org/10.1111/lang.12077

Stæhr, Lars S. 2008. Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal* 36: 139–152. https://doi.org/10.1080/09571730802389975

Tosi, Arturo. 2001. *Language and Society in a Changing Italy*. Clevedon: Multilingual Matters.

Zhang, Songshan & Xian Zhang. 2020. The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research* 26(4): 696–725. https://doi.org/10.1177/1362168820913998