Language Resources and Evaluation
https://doi.org/10.1007/s10579-024-09722-2

ORIGINAL PAPER

DILLo: an Italian lexical database for speech-language pathologists

Federica Beccaria[1] · Angela Cristiano[2] · Flavio Pisciotta[3] · Noemi Usardi[4] · Elisa Borgogni[5,6] · Filippo Prayer Galletti[4] · Giulia Corsi[5] · Lorenzo Gregori[6] · Gloria Gagliardi[7]

Accepted: 11 January 2024

Abstract

A novel lexical resource for treating speech impairments from childhood to senility: DILLo—*Database Italiano del Lessico per Logopedisti* (i.e., Italian Database for Speech-Language Pathologists) is presented. DILLo is a free online web application that allows extraction of filtered wordlists for flexible rehabilitative purposes. Its major aim is to provide Italian speech-language pathologists (SLPs) with a resource that takes advantage of Information and Communication Technologies for language in a healthcare setting. DILLo's design adopts an integrated approach that envisages fruitful cooperation between clinical and linguistic professionals. The 7690 Italian words in the database have been selected based on phonological, phonotactic, and morphological properties, and their frequency of use. These linguistic features are encoded in the tool, which includes the orthographic and phonological transcriptions, and the phonotactic structure of each word. Moreover, most of the entries are associated with their respective ARASAAC pictogram, providing an additional and inclusive tool for treating speech impairments. The user-friendly interface is structured to allow for different and adaptable search options. DILLo allows Speech-Language Pathologists (SLPs) to obtain a rich, tailored, and varied selection of suitable linguistic stimuli. It can be used to customize the treatment of many impairments, e.g., Speech Sound Disorders, Childhood Apraxia of Speech, Specific Learning Disabilities, aphasia, dysarthria, dysphonia, and the auditory training that follows cochlear implantations.

Keywords Lexical database · Speech impairments · Speech sound disorders · ICT for language · Clinical linguistics

Federica Beccaria, Angela Cristiano and Flavio Pisciotta have equally contributed to this work.

Extended author information available on the last page of the article

Published online: 23 March 2024

1 DILLo—*Database Italiano del Lessico per Logopedisti* ('Italian lexical database for speech pathologists')

DILLo—*Database Italiano del Lessico per Logopedisti* ('Italian Lexical Database for Speech-Language Pathologists') is an online resource created through collaboration between linguists, speech pathologists, and computer scientists. It is specifically intended to facilitate the development of speech and language therapy treatment tasks. Its main goal is to provide clinicians with word lists selected based on phonological, phonotactic, and morphological properties, and their frequencies. These cues are crucial for administering appropriate stimuli to the patients based on their rehabilitation needs. Therefore, DILLo is a lexical database that allows Speech-Language Pathologists (SLPs) to obtain a rich, tailored, and varied selection of suitable stimuli through a quick and easy search.

DILLo's applications include the treatment of Developmental Communication Disorders (specifically, organic or functional SSDs—Speech Sound Disorders, such as articulation disorders, phonological disorders, and CAS—Childhood Apraxia of Speech; cf. APA, 2013; Bowen, 2023), SpLDs—Specific Learning Disabilities (i.e., dyslexia, dysgraphia, and dysorthography), aphasia, dysarthria, dysphonia, and the auditory training that follows cochlear implantations.

Similar resources, intended to collect words coded for (psycho)linguistic properties, already exist for other languages. Examples are the MRC Psycholinguistic Database[1] (Coltheart, 1981; Wilson, 1988) for English and Norwegian Words[2] (Lind et al., 2015), or the Croatian Psycholinguistic Database[3] (Peti-Stantić et al., 2021). Nonetheless, analogous databases cannot automatically be transferred to other languages, especially if their targeted users and scopes are different. Thus, DILLo fills a gap in the context of Italian clinical linguistics and enriches the tools for speech-language treatments.

DILLo includes 7690 Italian words orthographically transcribed in Standard Italian, which may be consulted using three different query options. Each word has been manually coded and verified by a team of trained linguists to ensure its accuracy.

This paper is organized as follows: Sect. 2 provides an overview of the linguistic information on which DILLo was designed and, as a result, included in the tool. Section 3 instead illustrates DILLo's interface and search options to explain how to consult the database. Finally, Sect. 4 presents future developments and research directions for the tool.

[1] https://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm

[2] http://tekstlab.uio.no/ordforradet/en

[3] https://doi.org/10.17234/megahr.2019.hpb

2 Coded features of the words in *DILLo*

Sections 2.1–2.5 describe and motivate the various linguistic and paralinguistic features that DILLo's words have been coded for: phonological transcription and phonotactic structure, word frequency, regional variants, parts of speech, and ARASAAC pictograms.

2.1 Phonological transcription and phonotactic structure

DILLo contains the phonological transcription and phonotactic structure of all included words, thus offering the possibility to search for specific phonemes, or words with different syllabic lengths and stress types.

The inclusion of a phonological transcription is mainly due to the common notion of some phonemes being more complex than others, consequently leading to different ease of production. Epiphenomena of this complexity are the earlier acquisition of simpler phonemes by children (Gayraud et al., 2018; Kager et al., 2004; Stoel-Gammon, 1985) and their better preservation after brain damage (Buchwald, 2009; Galluzzi et al., 2015; Marquardt et al., 1979; Romani & Galluzzi, 2005; Wolk, 1986).

The articulatory complexity of Italian phonemes—and thus their relevance for SSDs—can be based on their acquisition trajectories. Phonemes acquired later in life can be considered more complex than those acquired earlier. For instance, the study by Zanobini et al. (2012) showed that children aged 36–42 months were able to produce voiceless plosives, bilabial and alveolar nasals (/m/ and /n/), and the alveolar lateral approximant /l/. More than 80% of the participants had an inventory encompassing at least 16 of the 23 consonants, as they successfully articulated voiced plosives (excluding /g/), all fricatives, the unvoiced postalveolar affricate /ʧ/, the alveolar trill /r/, and the approximants /j/ and /w/. A minority of the children exhibited production of the voiced postalveolar affricate /ʤ/, the voiced plosive /g/, the voiceless postalveolar fricative /ʃ/, the lateral palatal /ʎ/, or the nasal /ɲ/. Notably, none of the children included the affricates /ts/ and /dz/ in their consonant inventories. Another indirect measure of phonological complexity can be derived from information regarding which sounds more frequently or intensively undergo simplification, as this need to simplify suggests they are more complex than the others. Both typical and atypical simplification strategies (and timing) have been extensively described in the literature (Chilosi et al., 2014; Sabbadini et al., 2000). Nevertheless, phonetic and phonological development are not independent, as they are strictly intertwined together with lexical development (Vihman, 2017; Zamuner & Thiessen, 2018). For specific information regarding the order of acquisition of phonemes and phonological processes in Italian children, and their relationship with lexical development, readers can refer to Bortolini (1995), Bortolini et al. (1996), Zmarich and Bonifacio (2004), Pinton et al. (2014), and Viterbori et al. (2018).

In addition to the number of acquired phonemes, clinicians must also consider how their realization varies depending on their position within the word (Bortolini

et al., 1995, 1996; Bortolini & Leonard, 1996), as context-dependent diversification of outcomes may indicate phonological impairments (Bortolini & Leonard, 1991a, 1991b; Pinton et al., 2014). Moreover, phonological reorganization and coarticulation phenomena are also relevant in the rehabilitation of adults with dysarthria or apraxia of speech (Hardcastle & Tjaden, 2008; Ziegler & Von Cramon, 1985).

Aside from words segmental features, the suprasegmental level also impacts language acquisition. Suprasegmental factors that may affect acquisition are:

- Phonotactic distributions: sequences of sounds, licit in a language, do not all occur with the same frequency. Thus, the number of inputs children receive for these sequences might differ. This aspect has also been considered influential for SSDs (Coady & Aslin, 2004; Zamuner et al., 2004).
- The syllabic structure (e.g., length and type of onset or coda, tautosyllabicity of segments, number of syllables in the word, alternation between weak and strong syllables). For instance, the relevance of word length for SSDs has been thoroughly proven (Dollaghan & Campbell, 1998; Weismer et al., 2000). The number of syllables is also important in acquired language disorders; the output buffer deficits are affected by the word length effect (Caramazza et al., 1986; Patterson, 1986).

Evidence of a connection between complexity in acquisition and SSDs can be found among speech errors or simplification processes, that are indicators of speech impairments (e.g., symptoms of apraxia of speech) (Bortolini, 1995; Bortolini & Leonard, 1991a; Galluzzi et al., 2015; Zanobini et al., 2012). Both complex phonemes and complex suprasegmental features play a role in these processes. Segments that appear later in the acquisition trajectories, such as fricatives and affricates, present difficulties for patients; weak syllable deletion, metathesis, epenthesis, diphthong reduction, or consonant/vowel harmony are all sensitive to the complexity of the phonotactic structure of the word.

In people with SSDs, these difficulties with both phoneme-related and suprasegmental aspects can also affect other linguistic abilities, e.g., the marking of verbal agreement: an example is described by Leonard et al. (1992) and Bortolini et al. (1996). In neurotypical conditions, Italian is a dominant penultimate stress pattern language. Penultimate-stressed words are the most frequent; the second most common type of primary stress is on the antepenultimate (Nespor & Bafile, 2008). The distribution of different kinds of stress in the Italian lexicon is reflected by the words in *DILLo*, as Table 1 shows:

Table 1 Distribution of words with different types of stress in the DILLo database

Position of stress	n. of words
unstressed word	133 (1.7%)
final syllable	181 (2.3%)
penultimate syllable	6415 (83.4%)
ante-penultimate syllable	953 (12.4%)
< ante-penultimate syllable	8 (0.1%)
Total	7690

Table 2 Distribution of words of different syllabic lengths in the DILLo database

n. of syllables	n. of words
1	136 (1.8%)
2	1607 (20.9%)
3	2674 (34.8%)
≥4	3273 (42.6%)
Total	7690

Consequently, Italian children with a speech impairment may tend to pronounce antepenultimate-stressed words atypically to avoid less frequent syllabic sequences such as strong–weak-weak—as in *cantano* ['kan.ta.no], 'they sing', where they might delete the final weak syllable. However, in doing so, they also lose morphological information: by pronouncing *canta* ['kan.ta], 'he/she sings', they change the person of the verb.

A clinical application where both the frequency of phonotactic patterns and the length of the words are taken into account is the non-word repetition task (NRT). The NRT is considered one of the leading tests of speech perception, phonological encoding and assembly, and articulation; in fact, it has been frequently used as a diagnostic tool for a variety of SSDs (Munson et al., 2005; Coady & Evans, 2008; cf. D'Amico, 2000 on Italian).

Due to the centrality of segmental and suprasegmental aspects for speech impairments, each word in *DILLo*'s database was phonologically transcribed in Standard Italian (cf. §2.3.) and coded for its phonotactic structure. Transcriptions were manually carried out by two linguists and then manually re-examined by three other linguists to reduce human errors. The same procedure was applied to the phonotactic structure, extracting information regarding the type of stress and syllabic length (cf. Table 2) of words in *DILLo*.

2.2 Word frequency

Language processing and acquisition have been shown to be sensitive to frequency at any level of representation (Bybee & Hopper, 2001; Diessel & Hilpert, 2013; Ellis, 2002). Psycholinguistic research has shown the pivotal role of frequency and distributional information in the acquisition of grammatical patterns (Goldberg, 2006; Tomasello, 2003) and word classes (Diessel, 2007), configuring a generalization process on statistical grounds. Focusing on the lexicon, besides the earlier acquisition of more frequent words by children (Goodman et al., 2008; Swingley & Humphrey, 2018), word frequency also influences the speed of lexical retrieval in typical children, children with SSDs, and adults (Jescheniak & Levelt, 1994; Newman & German, 2002). This effect could be due to the repetitive use of a word, which makes it more entrenched, i.e., more easily and automatically accessible for the speaker, as shown by cognitive linguistics research (Bybee, 2007; Gries, 1999).

As expected from this brief overview, word frequency is relevant in treating language-impaired patients. For example, research on children with phonological delays has revealed that treating an erred sound in a high-frequency word enables a systemwide improvement of their phonological accuracy, thus affecting non-treated sounds (Gierut & Morrisette, 2012). Furthermore, Core Vocabulary Therapy, a treatment based on lists of highly frequent words, has proved effective in children with phonological planning deficits (Crosbie et al., 2005). Word frequency is a crucial parameter in lexical retrieval since high-frequency words have lower recognition thresholds (Morton, 1969). The main treatments for phonological and lexical deficits are based on this assumption (Basso, 2005; Marangolo, 2012). Because of the significant role of frequency in language acquisition, processing, and therapy, it was given a central function in the design of *DILLo*.

Word frequencies are treated according to the *Nuovo Vocabolario di Base dell'Italiano* (NVdB, i.e., 'New Basic Vocabulary of Italian') (De Mauro, 1999, 2016), that is based on frequency counts from an 18-million-word corpus of spoken and written Italian and on psycholinguistic experiments (Chiari & De Mauro, 2014). It divides the lexicon into three categories (De Mauro, 2003):

- Fundamental vocabulary (FO—*Vocabolario Fondamentale*): The 2000 most frequent lexemes in the Italian language, covering 86% of the corpus tokens (e.g., albero 'tree', avere 'to have', essere 'to be', mangiare 'to eat', felice 'happy', e 'and', o 'or').
- High-usage vocabulary (AU—*Alto Uso*): An additional 3000 frequently used lexemes, accounting for another 6% of the corpus tokens (e.g., aeroporto 'airport', biscotto 'cookie', cucire 'to sew', amaro 'bitter').
- High-availability vocabulary (AD—*Alta Disponibilità*): 2500 lexemes, less frequent than those in the two previous categories, yet essential in the everyday language of speakers (e.g., alluce 'big toe', peperoncino 'chili pepper', sganciare 'unhook', offeso 'hurt', prepotente 'overbearing').

Around 450 additional words were added, including frequent inflected forms of some of the lexemes. Currently, DILLo contains 7651 words[4] classified with De Mauro's frequency labels plus five additional categories. Table 3 shows the categories used and their distribution in the database.

Additionally, to facilitate the query of the database, it is possible to filter the query according to three levels of word frequency: *Alta* ('high'), *Media* ('medium'), and *Bassa* ('low'). This frequency-based subdivision of the words was carried out by coding them for their rank, according to the frequency list of the spoken corpus LIP available in De Mauro et al. (1993). Approximately, the level "high" includes the words ranked 1–3000 ($n = 2602$), the level "medium" those ranked 3001–20 000 ($n = 3787$), and the level "low" those from 20 000 onwards ($n = 926$).

[4] About 99% of the whole set of 7690 lemmas.

Table 3 Labels used to code word frequency in the DILLo database

Frequency labels	n. of words
AU-high usage vocabulary	2946 (38.6%)
AD-highly available vocabulary	2134 (28%)
FO-fundamental vocabulary	2047 (26.8%)
CO-common usage vocabulary	454 (5.9%)
TS-technical vocabulary	27 (0.4%)
ES-loanwords	16 (0.2%)
BU-low usage vocabulary	8 (0.1%)
RE-regional Italian vocabulary	2 (<0.1%)

2.3 Regional variants

DILLo also features a list of regional phonetic variants for each lexical entry. The Italian language situation can be characterized as a diaglossic *continuum* between Italo-Romance dialects, regiolects (regional Italians), and Standard Italian (Auer, 2005; Berruto, 2006). In particular, regiolects identifying features mainly come from local dialects (Grassi et al., 1997).

Regional Italians are the varieties currently employed in spoken communication across the peninsula (Poggi Salani, 2010). Standard Italian, instead, is a linguistic abstraction with no native speakers (Cerruti et al., 2017). Its phonetic and phonological features are not taught in schools, but they have nonetheless been codified. Several pronunciation models have succeeded over the years (Crocco, 2017); however, all these standard models are only followed by voice professionals, while the rest of the population uses pronunciations with different degrees of regionality and idiosyncrasy.

Therefore, including regional variants seemed a necessary step in creating a realistic and useful linguistic resource. Speech pathologists will likely encounter patients adhering to different regional standards, with variable rates of stigmatization, whose pronunciations could be different from those described by Standard Italian phonological transcriptions (cf. §2.1.). These regional features should not be regarded as speech errors or deviant pronunciations, as they are the expression of natural variation within language that ought to be normalized rather than stigmatized. This fact must be considered to avoid over- or under-identification of language impairment and confusion between difference and disorder (Ball & Bernhardt, 2008; Clark et al., 2020; Easton & Verdon, 2021). Instead, "[a]n unbiased approach towards working with culturally and linguistically diverse clients is imperative in speech pathology practice" (Clark et al., 2020). DILLo database includes regional variants to help speech pathologists discard inflectional differences and focus on pathological aspects.

The realization of the list of regional variants was carried out by a team of three linguists. The first step was the consultation of resources describing the phonological features of Italo-Romance dialects and regiolects. For this purpose, the

following works were used: Canepari (1980, 1992, 1999), Loporcaro (2009), and De Blasi (2014). By crossing data from these sources, the team looked for horizontal processes of convergence among regiolects (Cerruti & Regis, 2015), i.e., for phonological features shared by different regiolects (despite them being independently emerged in each variety or spread due to contact – cf. *koinè*: Siegel, 2001; Trudgill, 2004). The set of most common and spread features of regional variation determined which phenomena were to be represented by the regional variants included in DILLo. The team tried to avoid being too specific with this task since the resource is addressed to clinicians from all parts of Italy. An excessive amount of characterization may have been too grounded in extremely local features, lacking adaptability to speakers of different communities, and too chaotic for easy consultation. The phenomena that have thus so far been considered in the regional variants are listed as follows:

- Variation between mid-open and mid-closed stressed vowels in open syllables (e.g., ['beːne] and ['bɛːne], 'well'; ['koːsa] and ['kɔːsa], 'thing').
- Variation between open and closed stressed diphthongs in open syllables (e.g., ['fjoːre] and ['fjɔːre], 'flower'; [kame'rjeːre] and [kame'rjɛːre], 'waiter'; ['woːmo] and ['wɔːmo], 'man').
- Variation between mid-open and mid-closed stressed vowels in syllables closed by non-geminate /s/ or by a nasal consonant (e.g., ['festa] and ['fɛsta], 'party'; [kom'mento] and [kom'mɛnto], 'comment').
- Variation between mid-open and mid-closed stressed vowels in the following suffixes: -ett*, -ezz*, -ott*, -ozz*[5] (e.g., [stan'tsetta] and [stan'tsɛtta], 'little room').
- Variation between voiced and voiceless intervocalic /s/ (e.g., ['kaːsa] and ['kaːza], 'home').

Dubious cases were individually checked on the Italian dictionary provided as a handbook by Canepari (1992), which includes, for almost every entry, information about its variants in the main regional standards.

Finally, the team has manually transcribed the regional variants of each database entry based on the list above of phenomena. We hope to be able, in the future, to expand the list of considered phenomena while still maintaining a balanced amount of regional characterization.

2.4 Part of speech

The words in the database were also coded for their grammatical category, i.e., the "Part of Speech"—PoS (e.g., noun, verb, adjective, conjunction, etc.). This property has been proven relevant in the lexical acquisition, development, and processing in typical children and children with SSDs (e.g., Conti-Ramsden & Jones, 1997; Hansen, 2016; Sheng & McGregor, 2010). Furthermore, aphasiological data have led researchers to postulate a lexical class effect that affects word production

[5] The * means any vowel except /u/.

Table 4 Distribution of words based on their corresponding parts of speech in the DILLo database

Part of speech	n. of words
noun	4144 (53.9%)
proper noun	22 (0.3%)
verb	1876 (24.4%)
adjective	1316 (17.1%)
adverb	205 (2.7%)
pronoun	42 (0.5%)
conjunction	32 (0.4%)
preposition	23 (0.3%)
interjection	21 (0.3%)
article	5 (0.1%)
phonosymbol	4 (0.1%)
Total	7690

and comprehension independently of the semantic system (Rapp & Caramazza, 1997; Voghera & Laudanna, 2003). Most studies on the topic deal with nouns and verbs, and less attention has been devoted to other word classes (Tribushinina & Dubinkina, 2012), probably due to class frequency. Research has shown that nouns and verbs are more frequently used in parental input than other open-word classes (e.g., adjectives and adverbs) (Sandhofer et al., 2000). Despite being less investigated, closed word classes (e.g., clitic pronouns and determiners) may still be of interest for research: some studies have found them particularly vulnerable in children with SSDs, suggesting weak mastery of syntax (Thordardottir & Namazi, 2007; Befi-Lopes et al., 2013. For a review of clitics acquisition patterns in Italian, see Caprin & Guasti, 2009; Suozzi & Gagliardi, 2022).

In *DILLo*, it is possible to filter the query by selecting one of the eleven parts of speech the words were coded for, which include both open and closed word classes. However, it must be pointed out that the choice of the parts of speech to consider and the identification of their members are not straightforward. The amount, nature, and boundaries of word classes are still an open debate in linguistic theory, as they are often theory-dependent (for a discussion on word classes in Italian, cf. Colombo & Graffi, 2017). For this reason, words were coded according to the part of speech specified in their entry in the NVdB (De Mauro, 2016; cf. §2.2) to maintain a theory-agnostic and consistent coding scheme. Table 4 shows the list of parts of speech included in the database and their corresponding number of words.

Finally, as Table 4 shows, a distinction between common and proper nouns was made. This feature can be relevant for specific needs in therapy since research on aphasic patients has suggested that the processing of proper and common nouns relies on distinct mechanisms (Yasuda et al., 2000). Proper nouns have been entered into the database mainly to facilitate the creation of some minimal pairs.

2.5 ARASAAC pictograms

A total of 3745 words in the database present their respective ARASAAC—Aragonese Portal of Augmentative and Alternative Communication[6] pictogram (Palao, 2013). ARASAAC is a free resource under the Creative Commons license, designed to represent a message through images of people, objects, or abstract ideas, without cultural, linguistic, or cognitive boundaries. The portal has become a point of reference for many countries, such as Italy, Spain, France, Brazil, Finland, Germany, and Belgium (Paolieri & Marful, 2018). The worldwide success of this resource has grown since its first implementation in 2007, and today it is equipped with a collection of more than 10,000 pictograms in twenty different languages, representing one of the most widely used symbolic systems for the AAC—Augmentative and Alternative Communication (AAC).[7] As a matter of fact, contrary to earlier claims, AAC proved to be a useful tool that does not hinder speech development but rather supports the development of communication skills (Cress & Marvin, 2003; cf. Romski et al., 2015, for a review on the topic). Early AAC interventions provide long-term benefits, and a multimodal strategy that combines existing oral language interventions with AAC may improve the quality of speech impairment treatment and language learning (Olive et al., 2006; Wright et al., 2013). Thus, pictograms may play a crucial role in the rehabilitation process, offering broader possibilities for linguistic and communicative comprehension for patients (Costantino et al., 2011). Additionally, this multimodal approach supports not only patients with severe communication disorders but also young children and illiterate individuals (Beukelman & Mirenda, 2013). The images also allow SLPs to easily create games and activities to stimulate target words or elicit the naming of certain items. These latter uses of ARASAAC images are part of common rehabilitation practice. For this reason, DILLo allows visualization of ARASAAC symbol(s) associated with words whenever a query is carried out. For example, Table 5 reports the pictograms of a set of words searched in the database as representatives of each part of speech.

Paolieri and Marful (2018) pointed out that the associations between words and pictograms in the ARASAAC database were made according to subjective criteria. This aspect could lead to a misunderstanding between the professional and the patient during a diagnostic or clinical treatment. Therefore, even though Paolieri and Marful's (2018) psychometric analyses have shown a high degree of validity and reliability of ARASAAC pictograms, we would like to emphasize the need for a very cautious use of this tool.

[6] https://www.arasaac.org

[7] Augmentative and Alternative Communication represents fruitful clinical practice research. The AAC studies (and, when necessary, attempts to compensate for) the temporary or permanent communication disabilities of people with impairments in the production and understanding of speech (Glennen & DeCoste, 1997; Mcnaughton & Bryen, 2007; ISAAC, 2015).

Table 5 Examples of ARASAAC pictograms associated with words of each part of speech considered in *DILLo*

Word	Part of Speech	ARASAAC Pictogram
cane ('dog')	noun	
Africa	proper noun	
alzarsi ('to stand up')	verb	
felice ('happy')	adjective	
sempre ('always')	adverb	
egli ('he')	pronoun	
ma ('but')	conjunction	
tra ('between')	preposition	
buongiorno ('good morning')	interjection	
una ('a(n)', f. s.)	article (ind., fem.)	
bee	onomatopoeia	

3 *DILLo*'s design and search options

DILLo is a web application composed of a relational database and a modern HTML5 front-end. The tool supports live updates whenever new linguistic material is made available. The web interface is user-friendly and has a responsive design to be used by any device. Currently, the tool provides three search options: phoneme-based, grapheme-based, and a minimal pair extractor (cf. Fig. 1).

The phoneme-based query enables users to search for segments that the patient should have either acquired or lost based on scientific descriptions of its specific

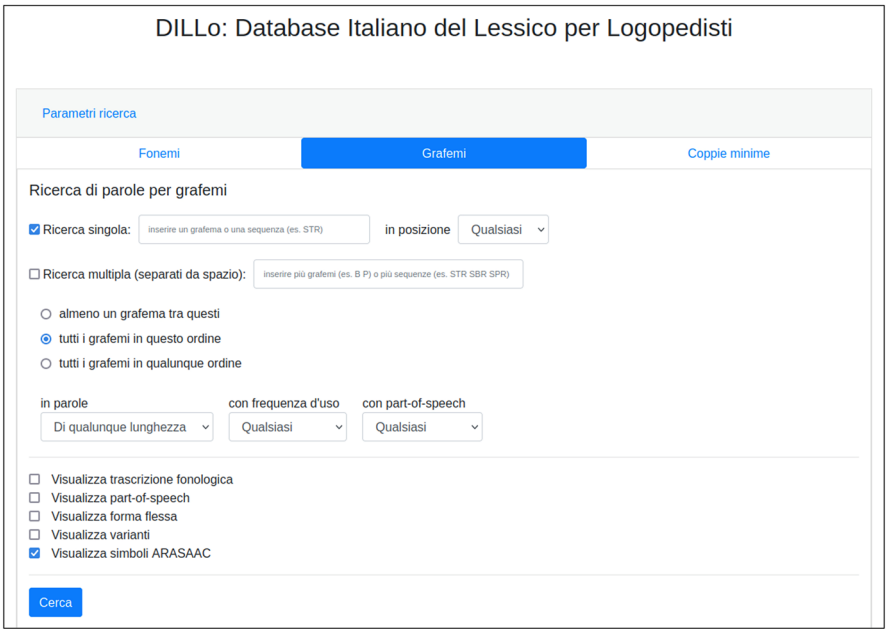

Fig. 1 DILLo search interface

disorder (e.g., developmental trajectories of phonological skills, articulatory complexity). The user is also able to select the position of the phoneme within the word (initial, medial, final) and to specify the phonological context (preceding and following phoneme), as clusters or specific types of vowels may have an impact on phoneme production.[8]

When searching by grapheme, it is possible to enter a grapheme (or more than one) or a string (or more than one) to find all the matching elements in the database; as in the phoneme-based research, the position in the word can be selected. Furthermore, in both search options, results can be filtered by choosing the number of syllables, the part of speech, and the frequency of the words (cf. Fig. 2).

The user can also extract two words that form a minimal pair so that the patient can be exposed to the phonemic relevance of the sound they are failing to use, discriminate or identify (Aimar et al., 2009; Barlow & Gierut, 2002; Gierut, 1989, 1991; Williams, 2000). This can be done by choosing the dedicated search option and selecting the contrasting phonemes; it is also possible to specify their position in the word.

For each element of the resulting word list, whichever search option has been used, the clinician can choose to show or hide several linguistic and non-linguistic information based on the features *DILLo*'s words have been coded for, namely:

[8] In fact, Bortolini and Leonard (1991a, p. 2) highlight how speakers affected by phonological disorders are sensitive to the "phonological details of the ambient language", other than the difficulty of specific articulatory gestures.

Tabella risultati

Lemma	Trascr. fono	PoS	Forme flessa	Varianti	Immagini
a	/a/	s.f.; s.m.inv.			
abachi	/ˈabaKi/	s.m.	pl.		
abbattimenti	/abbattiˈmenti/	s.m.	pl.	/abbattiˈmɛntI/	
abbonamenti	/abbonaˈmenti/	s.m.	pl.	/abbonaˈmɛnti/	
abbordo	/abˈbordo/	s.m.			
abitazione	/abitatˈʦjone/	s.f.		/abitatˈʦjɔne/	
abuso	/aˈbuzo/	s.m.		/aˈbuso/	
accademia	/akkaˈdɛmja/	s.f.		/akkaˈdɛmja/	

Fig. 2 DILLo results page

its phonological transcription, part of speech, regional variant(s) (if present), ARASAAC pictogram(s) (if present), and grammatical information associated with the word (e.g., number, gender, tense, person) if the word is an inflected form of a lexeme.

4 Future directions and conclusion

In summary, the DILLo database can be exploited to retrieve lexical units by SLTs with peculiar characteristics, empowering a customized logopaedic intervention. It allows SLTs to support the development or rehabilitation of specific language competences, such as lexical and phonological skills, which, as part of the overall rehabilitative approach, may contribute to a comprehensive improvement in language difficulties. Nevertheless, it is crucial to note that the tool may not fully address the complex needs of therapists, as patient difficulties often extend beyond the realm of isolated words.

The resource is in continuous development to improve the quality of the service provided to SLPs during their professional activity. In 2021, the tool underwent a two-month testing period by 13 Italian SLPs for the rehabilitation of clinically relevant cases (i.e., DLD – Developmental Language Disorder, SSDs, CAS, SpLDs, aphasia, dysarthria, dysphonia, auditory training), achieving an excellent satisfaction level with 92.30% of positive feedback (Usardi, 2021). However, the resource will

welcome further advice from the professionals using it to enhance as many functions as required.

Interesting future developments, some of which have successfully been included in lexical resources from other languages, may include:

- The addition of a query option to filter phonemes according to the articulatory movements required for their production (Hoch et al., 1986).
- The addition of an interrogation option allowing the extraction of "pseudo-minimal pairs" based on consonant length (e.g., ['ka:ne] 'dog' ~ ['kanne] 'reeds', ['fa:to] 'destiny' ~ ['fatto] 'fact'/ 'made') and lexical accent position (e.g., ['aŋkora] 'anchor' ~ [aŋ'ko:ra] 'again'/ 'still', ['pa:pa] 'pope' ~ [pa'pa] 'dad'). As a matter of fact, these phonological phenomena play a significant role in Italian. Specifically, consonantal quantity is contrastive on a regular basis and, as a result, many words may differ by only one geminated consonant (for a comprehensive review, the reader may refer to Bertinetto, 1981; Loporcaro, 1992 and Di Benedetto et al., 2021). Moreover, this feature holds significant typological relevance (e.g., compared to English, Spanish, and Chinese). Therefore, as an illustration, discrimination and production tasks can be administered to bilingual speakers with a first language different from Italian.
- The implementation of a virtual phonetic keyboard to facilitate entering IPA characters.
- The extension of the phonological context, in the phoneme-based query, to ± 2 phonemes adjacent to the selected one. For example, this option would allow comparing clusters based on the vowel following them (e.g., *tra* vs. *tro*).
- The possibility for the user to visualize the phonotactic structure of the word and its phonological transcription to assess syllable boundaries.
- The integration of other psycholinguistic information regarding semantic and lexical variables (Barca et al., 2002; Montefinese et al., 2014; Navarrete et al., 2019; Repetto et al., 2023; Spataro et al., 2019; Vergallito et al., 2020), i.e., the age of acquisition (Caselli et al., 2007; Marconi et al., 1993; Montefinese et al., 2019) and imageability of the word, or the frequency of the word in written production, both during childhood (Marconi et al., 1993) and adulthood (Bertinetto et al., 2005).
- The possibility to print a customized table with selected words and pictograms.
- The English translation of the browser interface to allow non-Italian researchers to access the resource.

In conclusion, DILLo represents a valuable resource for treating speech impairments, as it is the fruitful result of a cooperative effort between linguists and speech pathologists. The contribution of linguists to the creation of language rehabilitation tools is crucial to increasing and enhancing the database by providing controlled linguistic materials based on theoretical and methodological frameworks. In parallel, the participation of speech pathologists is necessary to ideate personalized treatments and design new features based on rehabilitation requirements.

Author contributions FB: data curation, writing; AC: data curation, writing; FP: data curation, writing; NU: data curation; EB: validation; FPG: software; GC: conceptualization; LG: conceptualization, software; GG: conceptualization, methodology, supervision.

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability The datasets generated during the current study are available from the corresponding author on request.

Declarations

Conflict of interest All authors declare that they have no conflicts of interest to disclose.

References

Aimar, E., Schindler, A., & Vernero, I. (2009). *Allenamento della percezione uditiva nei bambini con impianto cocleare*. Springer Science & Business Media.

APA-American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Publishing.

Auer, P. (2005). Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In N. Delbecque, J. van der Auwera, & D. Geeraerts (Eds.), *Perspectives on variation: Sociolinguistic, historical.* Comparative Mouton De Gruyter. https://doi.org/10.1515/9783110909579.7

Ball, J., & Bernhardt, B. M. (2008). First nations english dialects in Canada: Implications for speech-language pathology. *Clinical Linguistics & Phonetics, 22*(8), 570–588. https://doi.org/10.1080/02699200802221620

Barca, L., Burani, C., & Arduino, L. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers, 34*, 424–434.

Barlow, J. A., & Gierut, J. A. (2002). Minimal pair approaches to phonological remediation. *Seminars in Speech and Language, 23*(1), 57–68. https://doi.org/10.1055/s-2002-24969

Basso, A. (2005). *Conoscere e rieducare l'afasia*. Il Pensiero Scientifico.

Befi-Lopes, D. M., Pedott, P. R., Bacchin, L. B., & Cáceres, A. M. (2013). Word class and silent pauses in spoken narratives of children with specific language impairment. *CoDAS, 25*(1), 64–69. https://doi.org/10.1590/S2317-17822013000100012

Berruto, G. (2006). *Sociolinguistica dell'italiano contemporaneo*. Carocci.

Bertinetto, P. M. (1981). *Strutture Prosodiche dell'Italiano*. Accademia della Crusca.

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M. (2005). *Corpus e Lessico di Frequenza Dell'Italiano Scritto* (*CoLFIS*). Retrieved December 8, 2023 From .https://www.istc.cnr.it/en/grouppage/colfis

Beukelman, D. R., & Mirenda, P. (2013). *Augmentative & alternative communication: Supporting children and adults with complex communication needs* (4th ed.). Paul Brookes Publishing Co.

Bortolini, U. (1995). Lo sviluppo fonologico. In G. Sabbadini (Ed.), *Manuale di Neuropsicologia dell'Età Evolutiva* (pp. 203–241). Zanichelli.

Bortolini, U., Bonifacio, S., Zmarich, C., & Fior, R. (1996). Caratteristiche fonetiche di soggetti a basso rischio neonatale a 18, 21 e 27 mesi. *Età Evolutiva, 52*, 30–42.

Bortolini, U., & Leonard, L. B. (1991a). The speech of phonologically disordered children acquiring Italian. *Clinical Linguistics & Phonetics, 5*(1), 1–12. https://doi.org/10.3109/02699209108985499

Bortolini, U., & Leonard, L. B. (1991b). Phonology and children with specific language impairment: Status of structural constraints in two languages. *Journal of Communication Disorders, 33*(2), 131–149.

Bortolini, U., & Leonard, L. B. (1996). Phonology and grammatical morphology in specific language impairment:Accounting for individual variation in English and Italian. *Applied Psycholinguistics*, *17*, 85–104.

Bortolini, U., Zmarich, C., Fior, R., & Bonifacio, S. (1995). Word-initial voicing in the productions of stops in normal and preterm Italian infants. *International Journal of Pediatric Otorhinolaryngology, 31*, 191–206. https://doi.org/10.1016/0165-5876(94)01091-B

Bowen, C. (2023). *Children's speech sound disorders*. John Wiley & Sons.

Buchwald, A. (2009). Minimizing and optimizing structure in phonology: Evidence from aphasia. *Lingua, 119*(10), 1380–1395. https://doi.org/10.1016/j.lingua.2007.11.015

Bybee, J. L. (2007). *Frequency of use and the organization of language*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.001.0001

Bybee, J. L., & Hopper, P. J. (2001). *Frequency and the emergence of linguistic structure*. John Benjamins Publishing Company.

Canepari, L. (1980). *Italiano standard e pronunce regionali*. CLEUP.

Canepari, L. (1992). *Manuale di pronuncia italiana*. Zanichelli.

Canepari, L. (1999). *MaPI - Manuale di Pronuncia Italiana* (2nd ed.). Zanichelli.

Caprin, C., & Guasti, M. T. (2009). The acquisition of morphosyntax in Italian: A cross-sectional study. *Applied Psycholinguistics, 30*, 23–52. https://doi.org/10.1017/S0142716408090024

Caramazza, A., Miceli, G., & Villa, G. (1986). The role of the (output) phonological buffer in reading, writing, and repetition. *Cognitive Neuropsychology, 3*(1), 37–76. https://doi.org/10.1080/02643298608252669

Caselli, M. C., Bello, A., Rinaldi, P., Stefanini, S., & Pasqualetti, P. (2007). *Il Primo Vocabolario del Bambino: Gesti, parole e frasi. Valori di riferimento fra 8 e 36 mesi delle Forme complete e delle Forme brevi del questionario MacArthur-Bates CDI*. Franco Angeli.

Cerruti, M., Crocco, C., & Marzo, S. (Eds.). (2017). On the development of a new standard norm in Italian. In *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian.* De Gruyter Mouton. https://doi.org/10.1515/9781614518839-001

Cerruti, M., & Regis, R. (2015). The interplay between dialect and standard: Evidence from Italo-Romance. In E. Torgersen, S. Hårstad, B. Mæhlum, & U. Røyneland (Eds.), *Language variation European perspectives V.* John Benjamins. https://doi.org/10.1075/silv.17.05cer

Chiari, I., & De Mauro, T. (2014). The new basic vocabulary of Italian as a linguistic resource. In R. Basili, A. Lenci, & B. Magnini (Eds.), *Proceedings of the first italian conference on computational linguistics CLiC-it 2014 & the fourth international workshop EVALITA 2014.* Pisa University Press. https://doi.org/10.12871/CLICIT2014123

Chilosi, A. M., Lorenzini, I., Cerri, B., & Cipriani, P. (2014). Disprassia verbale evolutiva: inquadramento clinico e diagnosi differenziale con il disturbo fonologico. In L. Marotta & M. C. Caselli (Eds.), *I disturbi del linguaggio: caratteristiche, valutazione e trattamento.* Erickson.

Clark, E. L., Easton, C., & Verdon, S. (2020). The impact of linguistic bias upon speech-language pathologists' attitudes towards non-standard dialects of english. *Clinical Linguistics & Phonetics, 35*(6), 542–559. https://doi.org/10.1080/02699206.2020.1803405

Coady, J. A., & Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology, 89*(3), 183–213. https://doi.org/10.1016/j.jecp.2004.07.004

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*(1), 1–40. https://doi.org/10.1080/13682820601116485

Colombo, A., & Graffi, G. (2017). *Capire la grammatica: il contributo della linguistica*. Carocci.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A, 33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Conti-Ramsden, G., & Jones, M. (1997). Verb use in specific language impairment. *Journal of Speech, Journal of Speech, Language, and Hearing Research: JSLHR, 40*(6), 1298–1313. https://doi.org/10.1044/jslhr.4006.1298

Costantino, M. A., Anastasia, S., Bergamaschi, E., Bernasconi, L., Bianchi, A., Biffi, D., Cavallo, E., Cimò, V. G., Dall'Olmo De, C. V., Filippis, Errani, L., Festa, F., Ivan, D., Lanzini, L., Marini, M., Pugliese, L., & Zappa, G. (2011). *Costruire libri e storie con la CAA*. Erickson.

Cress, C. J., & Marvin, C. A. (2003). Common questions about AAC services in early intervention. *Augmentative and Alternative Communication, 19*, 254–272. https://doi.org/10.1080/07434610310001598242

Crocco, C. (2017). Everyone has an accent. Standard Italian and regional pronunciation. In M. Cerruti, C. Crocco, & S. Marzo (Eds.), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian.* De Gruyter Mouton.

Crosbie, S., Holm, A., & Dodd, B. (2005). Intervention for children with severe speech disorder: A comparison of two approaches. *International Journal of Language & Communication Disorders, 40*(4), 467–491. https://doi.org/10.1080/13682820500126049

D'Amico, A. (2000). Lo span di memoria verbale e la ripetizione di parole e non parole. Confronto tra buoni e cattivi lettori. *Rassegna Di Psicologia, 17*, 51–72.

De Blasi, N. (2014). *Geografia e storia dell'italiano regionale*. Il Mulino.

De Mauro, T. (1999). *Grande dizionario italiano dell'uso*. UTET.

De Mauro, T. (2003). *Guida all'uso delle parole. Parlare e scrivere semplice e preciso per farsi capire*. Editori Riuniti.

De Mauro, T. (2016). *Il Nuovo Vocabolario di Base della Lingua Italiana*. Internazionale. https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana

De Mauro, T., Mancini, F., Vedovelli, M., & Voghera, M. (1993). *Lessico di frequenza dell'italiano parlato*. ETAS libri.

Di Benedetto, M. G., Shattuck-Hufnagel, S., De Nardis, L., Budoni, S., Arango, J., Chan, I., & DeCaprio, A. (2021). Lexical and syntactic gemination in Italian consonants does a geminate italian consonant. Consist of a repeated or a strengthened consonant? *The Journal of the Acoustical Society of America, 149*(5), 3375. https://doi.org/10.1121/10.0004987

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology, 25*(2), 108–127. https://doi.org/10.1016/j.newideapsych.2007.02.002

Diessel, H., & Hilpert, M. (2013). Frequency effects in grammar. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics.* Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.120

Dollaghan, C. A., & Campbell, T. E. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136–1146. https://doi.org/10.1044/jslhr.4105.1136

Easton, C., & Verdon, S. (2021). The Influence of linguistic bias upon speech-language pathologists' attitudes toward clinical scenarios involving nonstandard dialects of english. *American Journal of Speech-Language Pathology, 30*(5), 1973–1989. https://doi.org/10.1044/2021_ajslp-20-00382

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143–188. https://doi.org/10.1017/S0272263102002024

Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia, 71*, 64–83. https://doi.org/10.1016/j.neuropsychologia.2015.03.007

Gayraud, F., Barkat-Defradas, M., Lahrouchi, M., & Ben Hamed, M. (2018). Development of phonetic complexity in arabic, berber, english, and french. *Canadian Journal of Linguistics / Revue Canadienne De Linguistique, 63*(4), 527–555. https://doi.org/10.1017/cnj.2018.9

Gierut, J. A. (1989). Maximal opposition approach to phonological treatment. *Journal of Speech and Hearing Disorders, 54*(1), 9–19. https://doi.org/10.1044/jshd.5401.09

Gierut, J. A. (1991). Homonymy in phonological change. *Clinical Linguistics and Phonetics, 5*(2), 119–137. https://doi.org/10.3109/02699209108985509

Gierut, J. A., & Morrisette, M. L. (2012). Age of word acquisition effects in treatment of children with phonological delays. *Applied Psycholinguistics, 33*(1), 121–144. https://doi.org/10.1017/S0142716411000294

Glennen, S., & DeCoste, D. C. (1997). *The handbook of augmentative and alternative communication*. Singular Publishing Group.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.001.0001

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35*(3), 515–531. https://doi.org/10.1017/S0305000907008641

Grassi, C., Sobrero, A., & Telmon, T. (1997). *Fondamenti di dialettologia italiana*. Laterza.

Gries, STh. (1999). Particle movement: A cognitive and functional approach. *Cognitive Linguistics, 10*(2), 105–145. https://doi.org/10.1515/cogl.1999.005

Hansen, P. (2016). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language, 37*(2), 205–225. https://doi.org/10.1177/0142723716679956

Hardcastle, B., & Tjaden, K. (2008). Coarticulation and speech impairment. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics.* Blackwell Publishing. https://doi.org/10.1002/9781444301007.ch32

Hoch, L., Golding-Kushner, K., Siegel-Sadewitz, V. L., & Shprintzen, R. J. (1986). Speech therapy. *In Seminars in Speech and Language, 7*(3), 313–326. https://doi.org/10.1055/s-2008-1064287

ISAAC Italy. (2015). Ruoli e responsabilità del Logopedista rispetto alla Comunicazione Aumentativa ed Alternativa: Relazione tecnica. *Argomenti di Comunicazione Aumentativa e Alternativa, 12*, 37–56.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Kager, R., Pater, J., & Zonneveld, W. (2004). *Fixing priorities: Constraints in phonological acquisition*. Cambridge University Press.

Leonard, L. B., Caselli, M. C., Bortolini, U., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition, 2*(2), 151–179. https://doi.org/10.1207/s15327817la0202_2

Lind, M., Simonsen, H. G., Hansen, P., Holm, E., & Mevik, B. H. (2015). Norwegian words: A lexical database for clinicians and researchers. *Clinical Linguistics & Phonetics, 29*(4), 276–290. https://doi.org/10.3109/02699206.2014.999952

Loporcaro, M. (1992). On the analysis of geminates in standard Italian and Italian dialects. In B. Hurch & R. A. Rhodes (Eds.), *Natural phonology.* The State of the Art Mouton de Guyter. https://doi.org/10.1515/9783110908992.153

Loporcaro, M. (2009). *Profilo linguistico dei dialetti italiani*. Laterza.

Marangolo, P. (2012). Riabilitazione dei deficit semantico-lessicali. In A. Mazzucchi (Ed.), *La riabilitazione neuropsicologica premesse teoriche e applicazioni cliniche*. Elsevier.

Marconi, L., Ott, M., Pesenti, E., Ratti, D., & Tavella, M. (1993). *Lessico Elementare. Dati statistici sull'italiano letto e scritto dai bambini delle elementari*. Zanichelli.

Marquardt, T. P., Reinhart, J. B., & Peterson, H. A. (1979). Markedness analysis of phonemic substitution errors in apraxia of speech. *Journal of Communication Disorders, 12*, 481–494. https://doi.org/10.1016/0021-9924(79)90011-X

Mcnaughton, D., & Bryen, D. N. (2007). AAC technologies to enhance participation and access to meaningful societal roles for adolescents and adults with developmental disabilities who require AAC. *Augmentative and Alternative Communication, 23*(3), 217–229. https://doi.org/10.1080/07434610701573856

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for english words (ANEW) for Italian. *Behavior Research Methods, 46*, 887–903. https://doi.org/10.3758/s13428-013-0405-3

Montefinese, M., Vinson, D., Vigliocco, G., & Ambrosini, E. (2019). Italian age of acquisition norms for a large set of words (ItAoA). *Frontiers in Psychology, 10*, 278. https://doi.org/10.3389/fpsyg.2019.00278

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*(2), 165.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*(5), 1033–1047. https://doi.org/10.1044/1092-4388(2005/072)

Navarrete, E., Arcara, G., Mondini, S., & Penolazzi, B. (2019). Italian norms and naming latencies for 357 high quality color images. *PLoS ONE, 14*(2), e0209524. https://doi.org/10.1371/journal.pone.0209524

Nespor, M., & Bafile, L. (2008). *I suoni del linguaggio*. Il Mulino.
Newman, R. S., & German, D. J. (2002). Effects of lexical factors on lexical access among typical language-learning children and children with word-finding difficulties. *Language and Speech, 45*(3), 285–317. https://doi.org/10.1177/00238309020450030401
Olive, M. L., Davis, T. N., Chan, J. M., Lang, R. B., & Dickson, S. M. (2006). The effects of enhanced milieu teaching and a voice output communication aid on the requesting of three children with autism. *Journal of Autism and Developmental Disorders, 37*, 1505–1513.
Palao, S. (2013). *ARASAAC Symbol Dictionary. Portal aragonés de la comunicación aumentativa y alternative*. Retrieved December 8, 2023.http://www.arasaac.org/
Paolieri, D., & Marful, A. (2018). Norms for a pictographic system: The Aragonese portal of augmentative/alternative communication (ARASAAC) system. *Frontiers in Psychology, 9*, 2538. https://doi.org/10.3389/fpsyg.2018.02538
Patterson, K. (1986). Lexical but nonsemantic spelling? *Cognitive Neuropsychology, 3*(3), 341–367. https://doi.org/10.1080/02643298608253363
Peti-Stantić, A., Anđel, M., Gnjidić, V., Keresteš, G., Ljubešić, N., Masnikosa, I., Tonković, M., Tušek, J., Willer-Gold, J., & Stanojević, M. M. (2021). The croatian psycholinguistic database: Estimates for 6000 nouns, verbs, adjectives and adverbs. *Behavior Research Methods, 53*(4), 1799–1816. https://doi.org/10.3758/s13428-020-01533-x
Pinton, A., Lena, L., & Zmarich, C. (2014). I disordini fonetico-fonologici. In L. Marotta & M. C. Caselli (Eds.), *I disturbi di Linguaggio Caratteristiche valutazione, trattamento.* ,.
Poggi Salani, T. (2010). Italiano regionale. In R. Simone (Ed.), *Enciclopedia dell'italiano I.* Istituto della Enciclopedia Italiana.
Rapp, B., & Caramazza, A. (1997). The modality-specific organization of grammatical categories: Evidence from impaired spoken and written sentence production. *Brain and Language, 56*(2), 248–286. https://doi.org/10.1006/brln.1997.1735
Repetto, C., Rodella, C., Conca, F., Santi, G. C., & Catricalà, E. (2023). The Italian sensorimotor norms: Perception and action strength measures for 959 words. *Behavior Research Methods, 55*, 4035–4047. https://doi.org/10.3758/s13428-022-02004-1
Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology, 22*(7), 817–850. https://doi.org/10.1080/02643290442000365
Romski, M., Sevcik, R. A., Barton-Hulsey, A., & Whitmore, A. S. (2015). Early intervention and AAC: What a difference 30 years makes. *Augmentative Alternative Communication, 31*, 181–202. https://doi.org/10.3109/07434618.2015.1064163
Sabbadini, L., De Cagno, A., Michelazzo, L., & Vaquer, M. (2000). *Il disordine fonologico nel bambino con disturbi del linguaggio*. Springer.
Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language, 27*(3), 561–585. https://doi.org/10.1017/s0305000900004256
Sheng, L., & McGregor, K. K. (2010). Object and action naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(6), 1704–1719. https://doi.org/10.1044/1092-4388(2010/09-0180)
Siegel, J. (2001). Koine formation and creole genesis. In N. Smith & T. Veenstra (Eds.), *Creolization and Contact.* John Benjamins,. https://doi.org/10.1075/cll.23.08sie
Spataro, P., Marques, V. R. S., & Rossi-Arnaud, C. (2019). Italian norms for the spontaneous completion of three-letter word stems: A preliminary study. *Current Psychology, 40*, 2221–2228. https://doi.org/10.1007/s12144-019-0160-4
Stoel-Gammon, C. (1985). Phonetic inventories, 15–24 months: A longitudinal study. *Journal of Speech Language and Hearing Research, 28*, 505–512. https://doi.org/10.1044/jshr.2804.505
Suozzi, A., & Gagliardi, G. (2022). The acquisition of the clitic ci among typically developing Italian preschoolers: preliminary data. *Linguistik Online, 116*(4), 77–114. https://doi.org/10.13092/lo.116.8891
Swingley, D., & Humphrey, C. (2018). Quantitative linguistic predictors of infants' learning of specific english words. *Child Development, 89*, 1247–1267.
Thordardottir, E. T., & Namazi, M. (2007). Specific language impairment in french-speaking children: Beyond grammatical morphology. *Journal of Speech, Language, and Hearing Research, 50*(3), 698–715. https://doi.org/10.1044/1092-4388(2007/049)

Tomasello, M. (2003). *Constructing a Language: A usage-based theory of language acquisition*. Harvard University Press.

Tribushinina, E., & Dubinkina, E. (2012). Adjective production by Russian-speaking children with specific language impairment. *Clinical Linguistics & Phonetics, 26*(6), 554–571. https://doi.org/10.3109/02699206.2012.666779

Trudgill, P. (2004). *New-dialect formation: The inevitability of colonial englishes*. Edinburgh University Press.

Usardi, N. (2021). *Ottimizzazione dell'applicativo DILLo "Database Italiano del Lessico per Logopedisti" e valutazione del suo utilizzo nella pratica clinica logopedica*. University of Florence.

Vergallito, A., Petilli, M. A., & Marelli, M. (2020). Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks. *Behavior Research Methods, 52*, 1599–1616. https://doi.org/10.3758/s13428-019-01337-8

Vihman, M. M. (2017). Learning words and learning sounds: Advances in language development. *British Journal of Psychology, 108*(1), 1–27. https://doi.org/10.1111/bjop.12207

Viterbori, P., Zanobini, M., & Cozzani, F. (2018). Phonological development in children with different lexical skills. *First Language, 38*(5), 538–559. https://doi.org/10.1177/0142723718784369

Voghera, M., & Laudanna, A. (2003). Proprietà categoriali e rappresentazione lessicale del verbo: Una prospettiva interdisciplinare. In M. Giacomo-Marcellesi & A. Rocchetti (Eds.), *Il verbo italiano. Studi diacronici, sincronici, contrativi, didattici Atti del XXXV Congresso Internazionale di Studi della Società di Linguistica Italiana (SLI).* Bulzoni. https://doi.org/10.1400/28229

Weismer, S. E., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43*(4), 865–878. https://doi.org/10.1044/jslhr.4304.865

Williams, A. L. (2000). Multiple oppositions: Theoretical foundations for an alternative contrastive intervention approach. *American Journal of Speech-Language Pathology, 9*(4), 282–288. https://doi.org/10.1044/1058-0360.0904.282

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers, 20*, 6–11. https://doi.org/10.3758/BF03202594

Wolk, L. (1986). Markedness analysis of consonant error productions in apraxia of speech. *Journal of Communication Disorders, 19*, 133–160. https://doi.org/10.1016/0021-9924(86)90016-X

Wright, C. A., Kaiser, A. P., Reikowsky, D. I., & Roberts, M. Y. (2013). Effects of a naturalistic sign intervention on expressive language of toddlers with down syndrome. *Journal of Speech, Language, and Hearing Research, 56*, 994–1008. https://doi.org/10.1044/1092-4388(2012/12-0060)

Yasuda, K., Beckmann, B., & Nakamura, T. (2000). Brain processing of proper names. *Aphasiology, 14*(11), 1067–1089. https://doi.org/10.1080/02687030050174638

Zamuner, T. S., Gerken, L., & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language, 31*(3), 515–536. https://doi.org/10.1017/s0305000904006233

Zamuner, T. S., & Thiessen, A. (2018). A phonological, lexical, and phonetic analysis of the new words that young children imitate. *Canadian Journal of Linguistics / Revue Canadienne De Linguistique, 63*(4), 609–632. https://doi.org/10.1017/cnj.2018.10

Zanobini, M., Viterbori, P., & Saraceno, F. (2012). Phonology and language development in Italian children: An analysis of production and accuracy. *Journal of Speech, Language, and Hearing Research, 55*(1), 16–31. https://doi.org/10.1044/1092-4388(2011/10-0228)

Ziegler, W., & Von Cramon, D. (1985). Anticipatory coarticulation in a patient with apraxia of speech. *Brain and Language, 26*(1), 117–130. https://doi.org/10.1016/0093-934X(85)90032-X

Zmarich, C., & Bonifacio, S. (2004). Gli inventari fonetici dai 18 ai 27 mesi d'età: uno studio longitudinale. In F. Albano Leoni, F. Cutugno, M. Pettorino, & R. Savy (Eds.), *Atti del Convegno Nazionale "Il Parlato italiano.* Il Parlato italiano, D'Auria Editore.

Authors and Affiliations

Federica Beccaria[1] · Angela Cristiano[2] · Flavio Pisciotta[3] · Noemi Usardi[4] · Elisa Borgogni[5,6] · Filippo Prayer Galletti[4] · Giulia Corsi[5] · Lorenzo Gregori[6] · Gloria Gagliardi[7]

✉ Gloria Gagliardi
gloria.gagliardi@unibo.it

1 Université Libre de Bruxelles, Brussels, Belgium

2 University of Groningen, Groningen, Netherlands

3 University of Salerno, Fisciano, Italy

4 Independent Researcher, Florence, Italy

5 Azienda AUSL Toscana Centro, Florence, Italy

6 University of Firenze, Florence, Italy

7 Department of Classical Philology and Italian Studies, University of Bologna, Via Zamboni, 32 – 40126, Bologna (BO), Italy