Automatically Estimating Textual and Phonemic Complexity for Cued Speech: How to See the Sounds from French Texts

Núria Gala, Brigitte Bigi, Marie Bauer
Aix Marseille Univ, Laboratoire Parole et Langage, CNRS
5 avenue Pasteur, 13100 Aix en Provence, France
nuria.gala@univ-amu.fr, brigitte.bigi@cnrs.fr, marie.bauer@etu.u-paris.fr

Abstract

In this position paper we present a methodology to automatically annotate French text for Cued Speech (CS), a communication system developed for people with hearing loss to complement speech reading at the phonetic level using hands. This visual communication mode uses handshapes in different placements near the face in combination with the mouth movements (called 'cues' or 'keys') to make the phonemes of spoken language look different from each other. CS is used to acquire skills in lip reading, in oral communication and for reading. Despite many studies demonstrating its benefits, there are few resources available for learning and practicing it, especially in French. We thus propose a methodology to phonemize written corpora so that each word is aligned with the corresponding CS key(s). This methodology is proposed as part of a wider project aimed at creating an augmented reality system displaying a virtual coding hand where the user will be able to choose a text upon its complexity for cueing.

Keywords: Cued Speech, hearing loss, phonetization, grapheme-phoneme annotation, textual complexity

1. Introduction

Linguistic resources for people with special needs (data-sets, annotated corpora, etc.) play an important role for building assistive AI applications. For people with hearing loss, there are a number of assistive technologies enabling speech perception such as listening devices or visual and sensory aids. There are also hearing aids involving speech recognition processing such as cochlear implants, etc. In addition, Cued Speech (CS) is a communication system that makes the language visible thus enhancing intelligibility of lip reading. In French, eight different handshapes in five different positions (cues) around the mouth are used to accompany natural speech to make the sounds of speech 'visible'. To date, however, there are very few resources for learning and practicing this code.

In this article, we propose a contribution for the French language: a methodology to automatically phonemize a written text and to estimate complexity features for learning and practicing CS.

The paper is organized as follows. In Section 2 we present CS and its mechanisms for visually conveying an oral language; an overview of existing language technologies for CS is also presented. Section 3 describes a methodology for automatically annotate a written text for cueing (using CS). Last section proposes an analysis of textual and phonemic complexity for CS in terms of readability and phonemization. We conclude with a discussion on the challenges of our project and we present future work, i.e. the creation of a French annotated corpus with complexity indicators that will be part of an augmented reality system displaying a virtual coding hand where the user will be able to choose a text upon its complexity for cueing.

2. Hearing loss, intelligibility of speech and Cued Speech

About 5% of the world's population live with disabling hearing loss (32 million are children) (WHO report, 2017). Failure to hearing impacts learning speech and its intelligibility. It also has an impact in learning to read: "while hearing children are learning to read words that they already have in their oral vocabulary, deaf children may be attempting to decode and understand words that they have never heard or seen before" (Reynolds, 2007). Consequently deaf adults read on average at only a 4th grade (ibid.), that is the estimated level at the end of elementary school.

Using CS enables the child to access to the reading mechanisms (phonemes-graphemes correspondences): phonemic awareness is increased, spelling and phonemic decoding are improved. The child can also get access to unlimited vocabulary, thus improving his/her lexical stock: "The lexicon developed by the deaf with Cued Speech has properties which are equivalent to the phonology of the hearing subjects (…) the internal representations of the words are compatible with their orthographic representations (…) this can prime the whole process of reading acquisition." (Alegria, 1990). The role of CS in learning to read is thus crucial, as shown by different studies (Ale-

LREC-COLING 2024, pages 1817–1824

gria, 1990; Leybaert et al., 2011; Trezek, 2017).

2.1. Origins and Aims of Cued Speech

Concerned with such low reading levels of most deaf population, CS emerged as a visual communication mode to complement speech reading at the phonetic level (Cornett, 1967). It is nowadays used as a tool for communicating in noisy environments and for learning to read an oral language in much the same way as a hearing child gets access to the different phonemes, thus allowing comprehension of the reading mechanisms.

Originally designed for American English, CS has been adapted to more than 65 languages[1]. In French it is known as *Langue française Parlée Complétée*, mainly learned through speech therapists and associations[2] like the ALPC[3] in France. Since 90 % to 95 % of deaf or hard of hearing children have hearing parents (Jones, 1989), CS is practiced by families willing to communicate with them with an oral language rather than with a sign language or in complement to.

The entire CS system is typically taught in several hours of classroom time. Although some parents explain that it is quite easy to learn, it demands practice. To date, however, there are not enough available materials to generalize such practice.

2.2. Cueing with Cued Speech

The production of speech naturally involves lip movements; for hearing people, both the acoustic information as well as the lipreading contribute to the phonological representation of sounds. Processing the audible acoustic information is influenced by the visual information (McGurk and MacDonald, 1976). For a better oral comprehension, every sound of a language should look different.

In French, for instance, the vowel /u/ is a height vowel compared to /a/ which is low. The jaw is high for a /u/ and low for a /a/; /i/ differs from /y/ by the roundedness of the lips, /i/ is unrounded while /y/ is rounded. However, in French - as in many other languages, many sounds look alike on the lips and can't be distinguished. To give an example, the words *bisou* and *minou*, 'kiss' and 'kitty': while the vowels are the same (/i/ and /u/), /b/ and /m/ are both bilabials and /n/ and /z/ alveolars. This makes it difficult to disambiguate both words only on the basis of lip reading. The notion of 'viseme' was introduced to refer to mutually confused phonemes that are deemed to form a single perceptual unit (Massaro, 1998).

In CS any syllable can be produced from both the handshapes representing consonants (C) and the hand position on the face representing vowels (V). A combination of a shape and a placement near the face is called 'cue' or 'key' (see Figure 1). Other syllabic structures can be produced with several keys - for example, a CCV syllable is cued with the two keys C then CV. Two different sounds, even if they look alike on the lips, are cued differently. Cued Speech keys match all the spoken phonemes, sounds are thus made visible, which results in a better understanding of speech.

2.3. Technologies for Cueing

While a large amount of work has been devoted to perception - e.g., (Nicholls and Mcgill, 1982; Alegria and Mattys, 1999; Bayard et al., 2019; Machart et al., 2020), production - e.g. (Leybaert et al., 2011), and CS synthesis - e.g., (Attina et al., 2004; Attina, 2005; Gibert et al., 2004; Wang et al., 2021), the resources for learning and training with CS remain scarce. This is especially true when it comes to available annotated (multimodal) corpora or training exercises. Such materials are of paramount importance considering that they could help deaf and hard of hearing or/and their families to learn the code itself (CS or any other code adapted to an oral language). They could also allow to practice and to improve the abilities in phoneme-grapheme correspondences.

2.4. Resources for Learning to Cue

Currently, CS can be mainly acquired through associations that provide materials and organize training courses for learning to cue, i.e. the National Cued Speech Association (NCSA)[4] in the US, or the *Association pour la Langue française Parlée Complétée* (ALPC) (see note 3) in France.

The traditional manner in teaching CS is by using highly illustrated books for practicing the keys corresponding to words and sentences. For the French CS, there are two booklets edited by the ALPC association: *Le petit clown 1 & 2* ('The little clown') by Valérie Sabbagh (2012). To become a fluent cuer it is important to practice, starting from one handshape and introducing new ones little by little, so that cueing becomes automatic.

In recent years, a number of videos and filmed materials are available through the Web, i.e. Youtube contains several CS channels proposing video lessons or showing how to cue specific words and expressions (hello, goodbye, etc.). Attempts have also been made to develop technologies with the goal to improve the learning and practice of CS based on 3D graphics (Arsov et al., 2010).

[1] International Academy Supporting Adaptations of Cued Speech (AISAC) `https://www.academieinternationale.org/`

[2] `https://cuedspeech.org/`

[3] `https://alpc.asso.fr/`

[4] `https://cuedspeech.org/learn/start-cueing/`

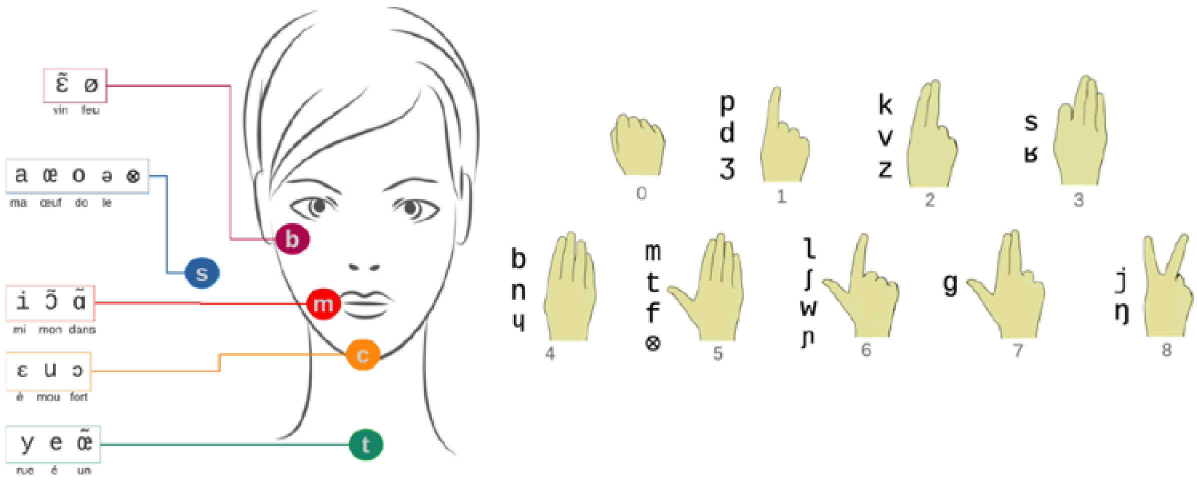

Figure 1: Cue chart for French CS

For French, there is also the CLeLfPC corpus, hosted by the French institutional repository Ortolang: a set of 4 hours of audio/video recordings, partly annotated (B. Bigi, 2022) at various levels including phonetics and CS-keys. The corpus is under the terms of the CC-by-NC-4.0 license, allowing to be used for any research on CS.

3. Automatic Annotation of Written Text for Cueing

In order to overcome the lack of resources for CS for French, the project *AutoCuedSpeech*[5] aims to develop an automatic generation system that will automatically cue. To train the system, and to provide learning materials to people learning to cue, we have developed a methodology to automatically annotate a written corpus in CS and to estimate textual and phonemic complexity for cueing. While the idea is to provide the corresponding key(s) to each word, the annotation cannot be done directly from the written text for several reasons. First, as Cued Speech is intended for speech, the text has to be phonetized. Second, a key in CS is a combination of a handshape for consonants and a placement in the face for vowels. Thus a key can only correspond to a vowel (V), a consonant (C) or a combination of consonant-vowel (CV), which means that a written syllable can sometimes be split into two or three keys, e.g. the word *triste* ('sad') is made of two syllables 'tris.te' but it is cued with a sequence of 4 keys 5s.3m.3s.5s corresponding to 't.ri.s.te'.
Numbers 1 to 8 correspond to a handshape, lowercase letters correspond to a placement near the face (s for side, m for mouth, t for throat, etc., see the cue chart for French, Figure 1). Last but not least, French is a highly opaque language for the graphemic-phonemic correspondences and there are different phenomena that add complexity to the transposition of the written text into CS (see section 4). Therefore, the method that we propose to automatically annotate a raw corpus and estimate its complexity requires normalization, manual liaisons annotation, and phonetization (grapheme to phoneme conversion, G2P).

3.1. SPPAS: automatic annotation and analysis of speech

SPPAS (Bigi, 2012, 2015) is a research software intended for corpus annotation. At present it proposes three major functionalities:

- Enrichment of a corpus with annotations that are produced automatically;
- Analysis of annotations;
- Conversion of annotated files to and from different file formats to ensure interoperability with manual annotation software (e.g. Praat, ELAN, ...).

3.1.1. Text Normalization

Some text transformations (Bigi, 2014) have to be performed by SPPAS for a correct phonemization: e.g., conversions of numbers (2 into *deux* 'two'), abbreviations (cm^2 into *centimètre carré* 'square centimètre'). This normalization also performs an automatic tokenization, converts all words into lowercase, remove punctuation, etc. The system takes as input a text and it produces a sequence of tokens, for instance:

input L'abat-jour a 1 an ! Parce qu'il est cassé...
output l' abat-jour a un an parce_qu' il est cassé[6]

[5] https://auto-cuedspeech.org/

[6] 'The lampshade is one year old because it is broken.'

3.1.2. G2P conversion

SPPAS enables to convert graphemes into phonemes (G2P), i.e. to automatically turn a normalized text into a new one with its constituent phones. The implemented approach is dictionary-based, with a solution to phonetize unknown words (Bigi, 2016). As the system is mainly dedicated to speech, one of the main consequences is that it proposes pronunciation variants instead of a unique standard phonemization. Speech variation can be the consequence of many different phenomena including the speaker accent, speech reductions, the liaison, etc.

In French, two phonetic phenomena have to be analyzed carefully in view to cueing: liaisons and glides. Liaison is a phenomenon where an orthographically-final consonant is mute except in certain environments, i.e. when it precedes a vowel, a mute h or a glide.

To our knowledge, an automatic tool for annotating French liaisons is not available, in part due to the arbitrary application of liaison rules: "Traditional accounts of liaison in French, mainly found in orthoepic textbooks, distinguish between liaisons that are termed obligatory or compulsory, those that are referred to as optional or variable, and those that are described as forbidden, erratic or impossible (...) How and when is liaison made? We are here in a delicate field, and there is no consensus to answer this question." (Boula De Mareüil et al., 2003).

The liaisons we decided to annotate are presented from the most frequent to the less in (Table 1). As for glides, there are three in French (semi-vowels or semi-consonants, Table 2). In CS they are considered as consonants and are cued at the beginning of a key.

Phoneme	Example	CS keys
z	vou**s** avez	vu.**za**.ve
t	ce**t** os	se.**to**.s
n	u**n** avion	œ̃ .**na**.v.jɔ̃
r	donne**r** un	do.ne.**r**œ̃
p	tro**p** important	t.ʁo.**p**ɛ̃.pɔ.ʁ.tɑ̃

Table 1: Examples of liaisons in French

Phoneme	Example	CS keys
j	fille	f.**je**
w	oui	**wi**
ɥ	fruit	f.ʁ.**ɥi**

Table 2: Examples of glides in French

When SPPAS is annotating speech, the answer to the question of speech variation is found into the audio recording. However, when working on written text, SPPAS is applying the principle of economy by choosing the shorter variant. For example, the French word *petit* ('small') can be pronounced both with or without a schwa /ə/ depending on the accent, and with or without a liaison depending on the context of use, which involves four different pronunciation variants. When using SPPAS as a G2P for written text, all the liaisons are then omitted; and so are the schwas.

The phonetization or not of the schwa has no consequence when coding with CS because the absence of vowel and the schwa are both at the same position (side). However, a solution had to be found for the liaisons. The problem was solved by adding explicitly each liaison into the text when it is required, for example the words sequence *vous avez un avion* ('you have a plane') is turned into *vous =z= avez un =n= avion* in order to be automatically phonetized into /vu zave zœ̃ navjɔ̃/.

3.1.3. CS coding annotation

SPPAS implements the automation of the production of keys (Bigi, 2023). Table 3 shows an example of the full automatic annotation process, from the raw text (orthography, line 1) to the CS keys (line 6)[7]. The input of the system is a text where liaisons have already been manually annotated (line 2). A blank is used to separate words after normalization; the symbol '-' separates phonemes and '.' the keys.

1	[orthography]	Vous avez un avion.
2	[input-text]	Vous =z= avez un =n= avion.
3	[normalized]	vous =z= avez un =n= avion
4	[phonetized]	v-u z a-v-e œ̃ n a-v-j-ɔ̃
5	[output-phon]	v-u.z-a.v-e.œ̃ .n-a.v.j-ɔ̃
6	[output-CS]	2-c.2-s.2-t.5-t.4-s.2-s.8-m

Table 3: Example of the annotation process

The phonetization produces a sequence of phonemes (line 4), which results in both a sequence of sounds (line 5) and their corresponding CS keys (line 6). Complex syllabic structures are split into several keys.

The automatic CS generator initially turns each phoneme into its class which is either labelled with C or V. An algorithm then specifies a sequence of handshape-position (C-V) pairs according to the rules of CS. This segmentation is performed thanks to a grammar implemented in a deterministic finite automata. Finally, when the sequence of class labels is segmented, the classes are turned back into phonemes. Each phoneme label is then mapped to its key code.

[7]The system can be tested on-line at https://auto-cuedspeech.org/annotate.html

4. Analyzing Textual and Phonemic Complexity for Cued Speech

The notion of complexity has already been addressed in the literature from different angles (for learning, for reading), considering different target populations and with a view of different applications. It encompasses a variety of factors that determine how challenging a given text is for a particular person and a for particular task (in general reading, i.e. decoding and/or comprehending). To give a few examples, a pilot study for French highlighted the effects of lexical complexity in dyslexic children (Gala and Ziegler, 2016). The authors showed that long words, complex syllable structures (several consonants) and a difference between the number of phonemes and graphemes are source of difficulty in reading (decoding) for children with dyslexia. In a recent study, (Yancey et al., 2021) proposed a readability model for French as a second language (L2) based on a deep learning model including linguistic, cognitive and pedagogical features.

With a view to cueing, our interest focused on textual and phonemic complexity, i.e. identifying the criteria that enable to decide a degree of difficulty for a text to be cued. While textual complexity can be measured with well-known features in the readability literature[8] (word length, text length, toke-type ratio or number or hapax, to name a few), the phonemic complexity of French has to be analyzed in detail specifically for cueing.

French is indeed a relatively consistent language from spelling to phonology (Ziegler et al., 1996) (*'a'* and *'à'* make the sound /a/; *'b'* corresponds to /b/, etc.). However it presents irregularities: the same grapheme can correspond to several phonemes, and the context is not helpful to disambiguate (e.g. the grapheme *'ch'* can correspond to the phoneme /ʃ/ in *choix* 'choice' or to a /k/ in *'echo'*). In addition, French present other phonological phenomena (silent letters, nasal vowels, liaisons, etc.) that have to be taken into account for estimating complexity.

4.1. Identifying Complexity Features

We explored different features to analyze the complexity of a French text with a view to cueing. They can be classified into two categories.

4.1.1. Readability features

Length has been a recurring proxy for readability since early works, i.e. (Flesch, 1948). "The variables in this group are the most studied and, despite their weaknesses, are still influential" (Wilkens et al., 2022):

1. Number of Words per Text (WT);
2. Number of Phonemes per Word (PHW);
3. Lexical diversity: number of Hapax[9] (HPX) or Type-tokens ratio[10] (TTR);
4. Lexical frequency[11](LFQ): calculated only for content words from the French resource Lexique3 (New et al., 2005).

4.1.2. Phonemic features

In addition to frequency-based information on the keys (see Table 4 and Table 5 next page), the identification of the phonemic features comes from the feedback of different French CS cuers from the ALPC association and from our practice as CS beginners.

1. CS keys frequency (KFQ): calculated on a reference corpus (see section 4.2);
2. Number of Consonant Clusters (CCL): when splitting a word and its syllables so that the configuration C, V or CV is always ensured, in some cases there are keys encoding single consonants;
3. Number of Liaisons (LSN): as liaisons are not visible in raw text, it is a feature that causes problems in human cueing when they practice CS from written text. The more a text presents liaisons, the higher the risk of errors;
4. Number of Glides or Semi-vowels (SVW): same than for liaisons and even higher risk of errors, because a human will tend to cue a vowel and not a consonant (see Table 2).

There can be other difficulties related to variations in pronunciation, e.g. making a difference between close/open vowels like for *rose* /ʁoz/ or /ʁɔz/ is a well-known regional variation. In our approach, the dictionary in SPPAS implements a standard pronunciation. Pronunciation varieties can be taken into account by modifying the dictionary. Note that a different vowel implies a different placement near the face and thus a different key (see Figure 1): /ʁoz/ 3**s**.2s (side) or /ʁɔz/ 3**m**.2s (mouth).

[8] Fabra (Wilkens et al., 2022) is a recent example of a toolkit for automatically computing textual characteristics for text readability in French.

[9] Hapax legomena are words that only occur once in a document, about 40-60% of a text.

[10] Variety of vocabulary which measures the ratio of unique words (types) to total words (tokens) found in a document.

[11] Lexical frequency is well known as a strong predictor of lexical complexity and readability.

4.2. Annotating a corpus in French

To implement our approach we have annotated an existing corpus for French, Alector (Gala et al., 2020) made of a collection of 100 original texts (50 narrative and 50 documentary), and their corresponding 100 adapted versions addressed to children 7-11 years old (2nd to 5th grades).

The corpus, with a total of 66 246 tokens and a mean length of 331 occurrences/text, was created with the aim of testing the effects of text simplification among different struggling readers.

Every text has two versions, the original one and a simplified variant where different transformations have been performed. Because the corpus is available on demand and some work on complexity has already been undertaken such as a classification by school grades, reading difficulty scores according to experimental data (Javourey-Drevet et al., 2022), etc. we decided to include it in our project.

The 200 texts from Alector were annotated with SPPAS and the readability and phonemic features were calculated. The total number of annotated keys is 91 826, with an average of 461 keys/text (see tables 4 and 5).

More than 25 000 occurrences in the corpus can be cued with a single key (e.g., frequent grammatical words such as determiners *le/la/un/une* ('the/a'), pronouns *il/elle/ils/elles* ('it/he/she/them'), conjonctions *que/mais/si/et/ou* ('that/but/if/and/or'), frequent prepositions *à/de* ('to/of'), etc.). Further analyses are currently underway to evaluate the annotations and to characterize them in terms of word lengths (number of phonemes).

Placement	Number of keys	Vowels
side (s)	36 194	a o œ ə ⊕
mouth (m)	19 407	ɑ̃ i ɔ̃
chin (c)	16 622	ɛ u ɔ
throat (t)	16 277	œ̃ e y
cheekbone (b)	3 286	ø ɛ̃

Table 4: Total number of keys by placement (vowels) in the CS annotated Alector corpus

Handshape	Number of keys	Consonants
5	22 626	m t f ⊕
3	18 669	s ʁ
6	14 835	l ʃ w ɲ
1	14 747	p d ʒ
2	10 978	k v z
4	5 949	b n ɥ
8	3 204	j ŋ
7	818	g

Table 5: Total number of keys by handshape (consonants) in the CS annotated Alector corpus

Tables 4 and 5 show the key's frequency distributions. As it was expected, the handshape with 5 deployed fingers (corresponding to sounds /m/, /t/, /f/ and no consonant ⊕) is the more frequent in combination with all the vowels. The more frequent placement is the side which enables to cue four vowels plus no vowel at all (for single consonants in consonant clusters, e.g. *fruit* in Table 2). Key 7 is the less frequent, which can be explained because it corresponds to the single phoneme /g/. Some further correlation studies are ongoing to better characterize the keys in terms of complexity, same for a complete descriptive analysis of the data. Nonetheless, with the frequency-based information on the keys, it is already possible to identify which is the corpus containing the highest number of a specific key or phoneme. Table 6 gives some examples. Not surprisingly, these are 5th grade texts (longer, with longer sentences and more lexical diversity). It is also possible to identify the longest corpus (id_188 with 950 keys) and the shortest one (id_14 from 2nd grade) with 131 keys.

5. Discussion and Future Work

In this position paper we have presented a methodology for automatically annotate raw text in French into keys for cueing and we implemented our approach on an existing French corpus. Cued Speech is a visual communication system enabling to 'see' the sounds of an oral language. To date, a number of hearing parents with deaf children use CS on a daily basis (at home, at school), yet the number of resources for learning and practicing the code is scarce.

Building resources for the deaf and hard of hearing and their families remains an open issue which requires collaboration among interdisciplinary researchers and field-work. By developing a CS encoding system, our contribution aims to meet the challenge of providing materials to learn and practice the code and to improve the abilities in phoneme-grapheme correspondences (skills of paramount importance for communication and for learning to read). Our project is thus innovative taking into account that the majority of the technology developed for CS is devoted to decoding. What's more, to date, there is no application enabling the user to choose a text based on a specific characteristic (type of key, combination of keys).

Within the AutoCuedSpeech project, the annotated corpus and the complexity features that we have identified will be part of an augmented reality system displaying a virtual coding hand where the user will be able to choose a text for training upon its complexity for cueing. The corpus and the videos will be available on a web platform by summer 2024.

Corpus id	Spécific Feature	Key(s)	Examples
id_188	more frequent key	6s	**bra**nche, oi**seau**, é**toi**le
id_200	more frequent keys with nasal V /ɛ̃/	3b/5b	**train**, che**min**, **main**
id_222	highest number of glides /j/	8	**scien**ce, spé**cia**lisé, trava**illent**

Table 6: Examples of corpus identified upon a specific complexity feature

Future work includes making the corpus available and creating more complete materials for CS. The system will also be tested among members of the ALPC association. Finally, a more detailed study on complexity for cueing from French text will be carried out, i.e. to integrate multimodal features such as the transitions between two keys taking into account their positions and handshapes.

6. Acknowledgements

This research was funded by the FIRAH (*Recherche Appliquée sur le Handicap*, Applied Disability Research), project APa2022_022[12]. The authors deeply thank the reviewers of the first version of the paper and the members of the French ALPC association, especially Maryvonne Zimmermann, for their valuable insights.

7. Bibliographical References

B. L. Alegria, J. Charlier and S. Mattys. 1999. The role of lip-reading and Cued Speech in the processing of phonological information in French-educated deaf children. *European Journal of Cognitive Psychology*, 11(4):451–472.

J. Alegria. 1990. Role played by Cued Speech in the identification of written words encountered by the first time by deaf children: a preliminar report. *Cued Speech Journal*, 4.

I. Arsov, B. Jovanova, M. Preda, and F. Preteux. 2010. On-line animation system for learning and practice cued speech. In *ICT Innovations 2009*, pages 315–325. Springer.

V. Attina. 2005. *La Langue Française Parlée Complétée: Production et Perception.* Phd thesis Ecole Doctorale Cognition, Langage, Education, Institut National Polytechnique de Grenoble - INPG.

V. Attina, D. Beautemps, M.-A. Cathiard, and M. Odisio. 2004. A pilot study of temporal organization in Cued Speech production of French syllables: rules for a Cued Speech synthesizer. *Speech Communication*, 44(1-4):197–214.

[12] https://www.firah.org/fr/appel-a-projets-general-2022.html

C. André B. Bigi, M. Zimmermann. 2022. CLeLfPC: a Large Open Multi-Speaker Corpus of French Cued Speech. In *Proceedings of the 13th International Conference on Language Resources and Evaluation (LREC'22)*, page 987–99, Marseille, France. European Language Resource Association (ELRA).

C. Bayard, L. Machart, A. Strauß, S. Gerber, V. Aubanel, and J-L. Schwartz. 2019. Cued Speech enhances speech-in-noise perception. *The Journal of Deaf Studies and Deaf Education*, 24(3):223–233.

B. Bigi. 2014. A multilingual text normalization approach. *Human Language Technology Challenges for Computer Science and Linguistics, LNAI 8387*, pages 515–526.

B. Bigi. 2015. SPPAS - multi-lingual approaches to the automatic annotation of speech. *The Phonetician*, 111-112:54–69.

B. Bigi. 2016. A phonetization approach for the forced-alignment task in SPPAS. *Human Language Technology. Challenges for Computer Science and Linguistics, LNAI 9561*, pages 515–526.

B. Bigi. 2023. An analysis of produced versus predicted French Cued Speech keys. In *Proceegins of the 10th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, ISBN: 978-83-232-4176-8*, pages 24–28, Poznań, Poland.

D. Bigi, B. Hirst. 2012. Speech phonetization alignment and syllabification SPPAS: a tool for the automatic analysis of speech prosody. In *Speech Prosody*, Shanghai (China).

P. Boula De Mareüil, M. Adda-Decker, and V. Gendner. 2003. Liaisons in French: a corpus-based study using morpho-syntactic information. In *Proc. of the 15th International Congress of Phonetic Sciences*.

R. O. Cornett. 1967. Cued Speech. *American annals of the deaf*, page 3–13.

R. Flesch. 1948. A new readability yardstick. *Journal of applied psychology*, 32(3).

N. Gala, A. Tack, L. Javourey-Drevet, and T. François. 2020. ALECTOR: A parallel corpus of simplified French texts with alignments of misreadings by poor and dyslexic readers. In *Language Resources and Evaluation Conference (LREC 2020)*, pages 1353–-1361, Marseille, France.

N. Gala and J.-C. Ziegler. 2016. Reducing lexical complexity as a tool to increase text accessibility for children with dyslexia. In *Computational Linguistics for Linguistic Complexity (CL4LC), workshop at COLING (Computational Linguistics conference)*, pages 59—66, Osaka, Japan.

G. Gibert, G. Bailly, F. Elisei, D. Beautemps, and R. Brun. 2004. Audiovisual text-to-cued speech synthesis. In *2004 12th European Signal Processing Conference*, pages 1007–1010. IEEE.

L. Javourey-Drevet, S. Dufau, T. François, N. Gala, J. Ginestié, and J.-C. Ziegler. 2022. Simplification of literary and scientific texts to improve reading fluency and comprehension in beginning readers of French. *Applied Psycholinguistics*, pages 1–28.

E. Jones. 1989. Evaluating the success of deaf parents. *American Annals of the Deaf*, 134(5):312–316.

J. Leybaert, M. Aparicio, and J. Alegria. 2011. The role of Cued Speech in Language Development of Deaf Children. *Oxford Handbooks Online. Deaf Studies, Language, and Education*, pages 1–16.

L. Machart, A. Vilain, H. Loevenbruck, and L. Ménard. 2020. Influence of French Cued Speech on consonant production in children with cochlear implants: an ultrasound study. In *ISSP 2020-12th International Seminar on Speech Production*.

D. W. Massaro. 1998. *Perceiving talking faces: from speech perception to a behavioral principle*. Cambridge, MA: MIT Press. Bradford Books.

H. McGurk and J. MacDonald. 1976. Hearing lips and seeing voices. *Nature*, 264(5588):746–748.

B. New, Ch. Pallier, and L Ferrand. 2005. Manuel de lexique 3. *Behavior Research Methods, Instruments, & Computers*, 36(3):516–524.

G. H. Nicholls and D. L. Mcgill. 1982. Cued Speech and the reception of spoken language. *Journal of Speech, Language, and Hearing Research*, 25 issue 2:262–269.

S. E. Reynolds. 2007. *An examination of Cued Speech as a tool for language, literacy and bilingualism for children who are deaf or hard of hearing*. Phd thesis program in audiology and communication sciences, Whashington University School of Medicine.

B. J. Trezek. 2017. Cued Speech and the development of reading in English: Examining the evidence. *The Journal of Deaf Studies and Deaf Education*, 22(4):349–364.

J. Wang, Z. Tang, X. Li, M. Yu, Q. Fang, and L. Liu. 2021. Cross-modal Knowledge Distillation Method for Automatic Cued Speech Recognition. *0.21437/Interspeech. Brno, Czech Republic WHO report*, pages 2986–2990.

R. Wilkens, D. Alfter, X. Wang, A. Pintard, A. Tack, K. P. Yancey, and T. François. 2022. FABRA: French aggregator-based readability assessment toolkit. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 1217–1233, Marseille, France. European Language Resources Association.

K. Yancey, A. Pintard, and T. Francois. 2021. Investigating readability of French as a foreign language with deep learning and cognitive and pedagogical features. *Lingue e linguaggio*, 20(2):229–258.

J.-C Ziegler, A. Jacobs, and G. Stone. 1996. Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers*, 28:504–515.