



**HAL**  
open science

## Development of an application to test young L2 learners' phonological awareness autonomously

Marie-Pierre Jouannaud, Emilie Magnat, Emilie Charles, Coralie Payre-Ficout, Mathieu Loiseau

### ► To cite this version:

Marie-Pierre Jouannaud, Emilie Magnat, Emilie Charles, Coralie Payre-Ficout, Mathieu Loiseau. Development of an application to test young L2 learners' phonological awareness autonomously. EuroCALL 2024: CALL for humanity, CALL for Humanity - EuroCALL 2024 Short Papers, Universitat Politècnica de València, pp.231-239, 2024, 978-84-1396-257-3. 10.4995/eurocall2024.2024.19074 . hal-04887372

**HAL Id: hal-04887372**

**<https://hal.science/hal-04887372v1>**

Submitted on 14 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.






L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

## Development of an application to test young L2 learners' phonological awareness autonomously

Marie-Pierre Jouannaud<sup>a</sup>, Emilie Magnat<sup>b</sup>, Emilie Charles<sup>c</sup>, Coralie Payre-Ficout<sup>d</sup> and Mathieu Loiseau<sup>e</sup>

<sup>a</sup>Université Paris 8, Transcrit, , email: [marie-pierre.jouannaud@univ-paris8.fr](mailto:marie-pierre.jouannaud@univ-paris8.fr); <sup>b</sup>Université Lyon 2, ICAR UMR 5191 CNRS, , email: [emilie.magnat@univ-lyon2.fr](mailto:emilie.magnat@univ-lyon2.fr); <sup>c</sup>Université Lyon 2, , email: [emilie.charles1@gmail.com](mailto:emilie.charles1@gmail.com); <sup>d</sup>Université Grenoble Alpes, LIDILEM, , email: [coralie.payre-ficout@univ-grenoble-alpes.fr](mailto:coralie.payre-ficout@univ-grenoble-alpes.fr) and <sup>e</sup>INSA-Lyon, LIRIS UMR CNRS 5205, , email: [mathieu.loiseau@liris.cnrs.fr](mailto:mathieu.loiseau@liris.cnrs.fr)

How to cite: Jouannaud, M.-P.; Magnat, E.; Charles, E.; Payre-Ficout, C.; Loiseau, M. (2024). Development of an application to test young L2 learners' phonological awareness autonomously. In Y. Choubsaz, P. Díez-Arcón, A. Gimeno-Sanz, J. Hriňák, X. Liashuk, S. Pokrivčáková & H. Vančová (Eds.), *CALL for Humanity - EUROCALL 2024 Short Papers*. <https://doi.org/10.4995/EuroCALL2024.2024.19074>

---

### Abstract

*Phonological Awareness (PA) plays a crucial role in L1 and L2 reading acquisition and L2 pronunciation, but its relation to L2 listening has been studied to a lesser degree. In order to study the relationship between the development of PA as a meta-skill and the use of a foreign language app targeting listening skills for young learners, a reliable test was needed. The goal of this study was to design a PA test in English and in French to be administered on tablets, taken by French children autonomously (without the need for one-on-one administration) and assessed automatically. This requires the tasks to be transformed into selected response tasks to be automatically scorable. In order for instructions to be comprehensible to young learners they must be carefully scripted, illustrated and contextualized (in our case, with a family of monsters enjoying words starting or ending with certain sounds). The application was tested with 65 intact classes of 1st graders (6-7 years old) learning English in France, whose schools participated in a larger project focused on L1 French reading fluency and L2 English listening comprehension. It proved to be usable and reliable. Further validity studies are planned.*

**Keywords:** *phonological awareness; young learners; English as a foreign language; language testing; tablet application.*

---

## 1. Introduction

Phonological Awareness (PA), or the ability to manipulate segments of spoken language, is one of the cognitive factors that determine the level of competence in both native (L1) and foreign (L2) languages. Given its importance, various PA tests have been developed in different languages. In this article, we describe the development of a PA test in two languages (English and French) to be administered on tablets, and to be taken by French first and second graders (six to seven years old) autonomously and assessed automatically.

This study is a follow-up to a larger project (Mandin et al., 2021) in which two tablet apps were developed to teach L1 French reading fluency and L2 English listening comprehension to French first grade schoolchildren.

After finding an unexpected effect of the L2 English app on L1 French phonological awareness, we decided to investigate the effect and develop our own PA testing tool, using tablet-based exercises to facilitate test-taking and data processing for future experiments.

## **2. Defining and testing phonological awareness in second language learning**

### **2.1 Definition of phonological awareness**

Phonological awareness has been the subject of much research since the late 1970s and early 1980s (for a review, see Gillon, 2018), and several definitions of phonological awareness have been proposed. According to Anthony and Francis (2005, p. 256), "phonological awareness refers to one's ability to recognize, discriminate, and manipulate the sounds in one's language, regardless of the size of the word unit that is the focus". It is a metacognitive skill in the sense that it requires being able to focus on the form (the sounds) of the linguistic material regardless of meaning, clearly distinguishing phonological awareness from normal communicative use of language. The term 'awareness' implies a certain degree of conscious control, i.e. it is not just a question of being able to discriminate between different sounds or minimal pairs (phoneme discrimination) but requires the ability to compare or create new (nonsense) forms regardless of their meaning (or absence of it). Language thus becomes an object of observation and reflection (White & Ranta, 2002).

The size of the phonological units that speakers are able to manipulate seems to depend mainly on maturity. Carroll et al.'s (2003) study with 67 preschool children indicates that children first develop their implicit sensitivity to large units of language (syllables and rhymes) before becoming sensitive to smaller units such as phonemes. Some authors defend the idea that these are separate skills, while others support the hypothesis of a continuum with a single skill that gradually evolves from the processing of the largest units, i.e. syllables, to the processing of the smallest units, i.e. phonemes (Anthony & Lonigan, 2004; Goswami & Bryant, 1990), via infra-syllabic units (onsets and rhymes). Phonemic awareness is initially more difficult to acquire than syllable and rhyme awareness, and develops as we learn to read (Høien et al., 1995; Liberman et al., 1974).

Phonological awareness is a key component for the different language skills. It has long been a focus of study in L1 reading and writing skill development, and phonemic awareness, in particular, has emerged as a key precursor of L1 reading acquisition (at least in alphabetical languages). It has also been implicated in L2 reading (Haigh et al., 2011) and L2 pronunciation (Kivistö de Souza, 2015), but its relation to L2 listening has been studied to a lesser degree (Li et al., 2013). More generally, as a cognitive variable, it has been shown to be correlated with the lexicon and morphosyntax for both L1 and L2 speakers (e.g. Hopp et al., 2019).

While some studies have looked at the links between phonological awareness in L1 and L2, Melby Lervåg and Lervåg's (2011) meta-analysis pointed out that "[t]he fact that we found significant correlations between L1 and L2 on [...] phonological awareness skills can be seen as evidence for cross-linguistic transfer. It should be noted, however, that the findings only demonstrate a relationship, and not the mechanism that causes it" (p.128). This is in line with the work of various researchers who consider phonological awareness not to be language-specific, but to be a general meta-linguistic skill (Comeau et al., 1999; Cummins, 1979). The results of the initial project (alluded to above) in which the use of an English listening app was found to increase PA in French seem to concur with this hypothesis.

### **2.2 Traditional phonological awareness tests**

Traditionally, phonological awareness tests focus on one language at a time, for example English (CTOPP), French (N-EEL, Chevrie-Muller & Plaza, 2001; BALE, Jacquier-Roux et al., 2010), or German (BAKO, mentioned in Kersten et al., 2024). Also, because these tests are often used for their predictive power of future reading skills they are calibrated with a native population. Although some researchers have used tests designed for native speakers as part of their research into L2 learning (Magnat, 2013), they acknowledge that this use with non-native populations is not optimal if learners know that the test is in a foreign language and if they lack the lexical knowledge required. Finally, these tests are often only commercially available and are not free or open access.

In terms of test task characteristics, as mentioned above, PA test tasks can vary as to the size of units being manipulated (in order of decreasing difficulty): phonemes, rhymes or syllables. The tasks themselves usually belong to the following categories:

- comparison tasks, comprising judgment tasks (e.g. do these two words rhyme, or does this word contain this sound?), similarity tasks (which word starts/ends with the same phoneme?), and oddity tasks (e.g. which word does not start/end with the same sound(s) as the others?);
- counting or positioning (how many syllables are there in this word, or where is this sound in this word?);
- manipulation tasks are the most commonly used (here, examples with phonemes): segmenting a word into phonemes, blending phonemes to make a word, deleting a phoneme from a word, substituting one phoneme for another, reversing the order of phonemes in a word.

From the test takers' point of view, judgement tasks can be answered with 'yes' or 'no', oddity and counting tasks with a number (the number of the odd one out for oddity tasks), but manipulation tasks require learners to produce spoken language, which is always more difficult for L2 speakers.

Most of the time, PA tests are administered by a trained therapist, teacher, researcher or assistant. Training is necessary, especially when the tests are normed (i.e. performance on the test can be interpreted by reference to a larger population, ALTE, 1998, p. 154), in order to ensure that scoring is consistent with instructions and that future decisions based on scores are valid (knowledge of the phonetic alphabet is also sometimes necessary). Administration requires one-on-one time with each child and funds to pay for the tests. Additionally, they are often pen-and-pencil tests (e.g. BALE for French), and require the answers to be transferred to a spreadsheet afterwards if group results are to be analyzed for research purposes. Computerized tests do exist: in French, for example, a test was developed by Sprenger-Charolles et al. (2005), but it is still administered by an adult one-on-one (its main advantages being automatic scoring and response time availability). We also do not know whether it is still available and are not aware of any recent studies having used it for research. Other initiatives have arisen to develop technology-based tools and proven that they could be as reliable as paper-based tests (Carson et al., 2014; Kiss & Csapó, 2024; Meira et al., 2023).

In summary, traditional tests of phonological awareness focus on the manipulation of phonological elements in a single language. In this sense, they are not consistent with one of the hypotheses about phonological awareness, namely that it is not language specific. It is also worth noting the limitations of conventional tests: they have to be administered by one person to each learner individually, which requires time and qualified human resources. Because of these drawbacks, we set out to design a PA test in two languages (English and French) to be administered on tablets, taken by children autonomously and assessed automatically. The development of the test and its large-scale field testing to make sure that it is usable and reliable will be described in the following sections.

### 3. Methods

#### 3.1 Test design features

The obvious choice for a technology-based test is to use tablets (Charles et al., 2022). Tablets are cheaper than computers and take up less space, making it easier to equip an entire class. Moreover, most children are familiar with tablets and are enthusiastic about using them.

In order for children to take the test autonomously, the instructions must be as short as possible, simple, and accompanied by several training items before scored test items can be administered. All instructions (provided here in French since it is the learners' L1) must be spoken to make sure that children who cannot yet read can understand them. Additionally, to ensure engagement and aid comprehension, it is useful to contextualize the tasks. To this end, a group of colorful monsters was created, called Monstruas or Monstruos, who like to play with sounds and need help to achieve various goals (usually eating their favorite foods). A screenshot of the interface is provided in Figure 1.

The second requirement is for tasks to be automatically scorable. This means that no oral answers are possible (unless automatic speech recognition were to be used, Wills et al., 2023), and that all items need to be transformed into selected response tasks, where answers are given by clicking on the correct option or by dragging it to a drop zone. Moreover, for some tasks, it is common in monolingual tests to have the participants rely on their lexicon and choose from images without the words being pronounced (Kiss & Csapó, 2024). Since the learners had little or no vocabulary in L2 English we told them that the items (in French and English) were non-words, thus inviting them to focus on manipulating phonological elements independently of meaning. This was modelled on part of the CTOPP (English) and N-EEL (French) tests. Non-words were represented visually by invented objects or food. Finally, in order to test the hypothesis that PA is a cross-linguistic skill, items in two languages were included (French and English).

Our test is thus composed of four subtests addressing the three levels of granularity (rhyme, syllable and phoneme) and using four types of PA tasks (oddity and similarity judgment, blending and segmentation). Each task contains between 16 and 24 items of increasing difficulty, alternating between French and English and male or female speakers:

1. In the syllable blending task (Figure 1) the child needs to select the item resulting of the blending of two syllables by touching one of the 3 symbols at the bottom of the screen.

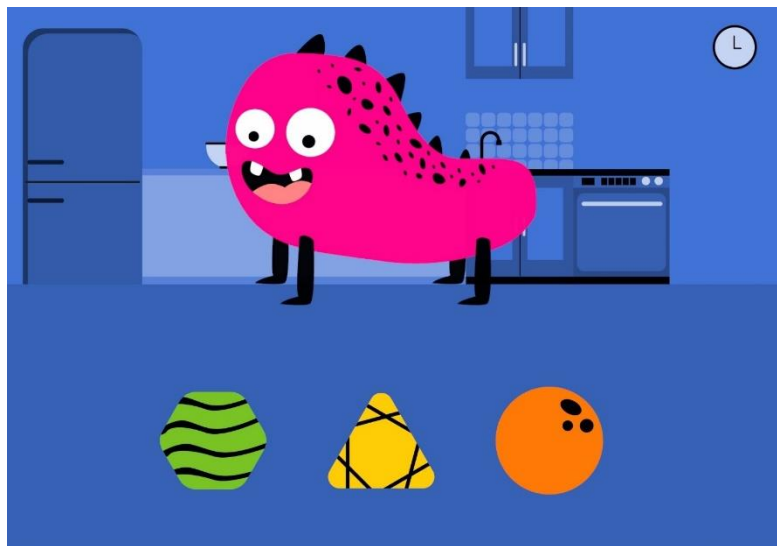


Figure 1. Screenshot of the interface for the syllable blending task.

2. In the phoneme segmentation task (Figure 2) the child first hears a non-word then touches the four symbols at the bottom of the screen to listen to four different phonemes, then drags each one to the correct position in order to reconstruct the non-word (one of the symbols is a distractor phoneme).

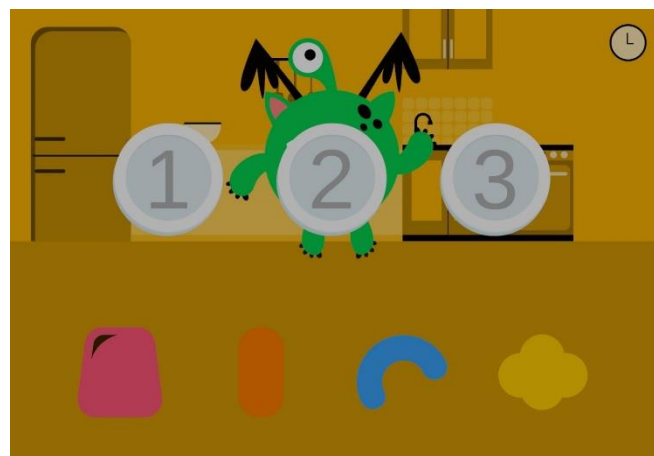


Figure 2. Screenshot of the interface for the phoneme segmentation task.

3. In the rhyme similarity task the children select the item (out of three) that rhymes with the name of the monster. The interface is similar to Figure 1 with a different monster and a different color background.

4. In the initial syllable oddity task (also with an interface similar to Figure 1) the goal is to select the item that does not start with the same syllable as the others.

As mentioned above, this test was part of the post-test for a research project. The various tasks were carried out in three sessions, with automatic blocking between sessions to prevent learners from doing several sessions in succession. In the first session the syllable blending task was taken immediately after an English listening comprehension test. The second session consisted of the phoneme segmentation task. The third session included the rhyme similarity task and the initial syllable oddity task.

### 3.2 Experimental design

The application, called ‘Evalulu’ was tested with 652 first graders (six to seven years old) learning English in France, whose schools participated in a larger project focused on L1 French reading fluency and L2 English listening comprehension. They belonged to 52 intact classes. The pretest for the larger experiment served as a pilot, after which the worst performing items and tasks were discarded for the post-test, mostly because they were too difficult (for example, there was initially a phoneme inversion task which was not kept). Some instructions were also simplified or reworded to make them easier for the children to understand. The pilot also enabled us to order items in order of ascending difficulty within each task. Here, we present the results of the post-test which took place at the end of the school year. All analyses were performed with *R*.

## 4. Results

### 4.1 Usability

All teachers were able to organize administration within their classes in the way they thought was most appropriate. Each class had between four and six tablets, and children took the test in small groups autonomously with headphones while the teacher worked with other students in the same classroom. It must be noted that the children (and their teachers) were familiar with the use of tablets since they had been using them for the duration of the school year.

All answers were collected automatically either during the test (in classrooms equipped with Wi-Fi), or after test administration by connecting the tablets to Wi-Fi. Because of log malfunction the time needed to complete the test is not known at this time.

### 4.2 Reliability and item characteristics (difficulty, discrimination)

As can be seen in Table 1, all tasks proved reliable (internally consistent), with Cronbach's alpha values above .7 (Laveault & Grégoire, 2014, p. 119) and reaching .91 for the segmentation sub-test. The range of item difficulty was also acceptable, with values between .2 and .8 (Bachman, 2004, p. 138), although the last two tasks were still difficult for this population (no item of syllable oddity or rhyme similarity judgment was successfully attempted by more than two thirds of our sample).

**Table 1.** Summary of test task characteristics: number of items, Cronbach's alpha, range of item difficulty, (numbers in parentheses correspond to discarded items), range of item discrimination, and mean score.

Id	Task	# items	$\alpha$	Item difficulty	Item discrim.	Mean score (sd)
1	Syllable blending	16	.78	.53 – .74	(.17) .28 – .43	10.5 (3.61)

2	Segmentation	24	.91	.30 – .77	.35 – .68	13.77 (6.64)
3	Rhyme similarity	16	.71	(.17) .35 – .66	(-.17) .18 – .50	8.11 (3.34)
4	Syllable oddity	16	.72	.34 – .58	(.15) .19 – .47	7.45 (3.43)

Test items also need to display adequate discrimination, i.e. higher-level students need to be more successful at them than lower-level students (the index used is a point-biserial coefficient, which is a correlation between item score and total test score excluding the item). The minimum acceptable value is considered to be .15 or .2 (Laveault & Grégoire, 2014, p. 211). Because some items had low discrimination values we discarded the lowest performing item in three sub-tests (syllable blending, rhyme similarity and syllable oddity), resulting in 15 items for these sub-tests. The segmentation sub-task, while being very different from a classical segmentation task requiring learners to produce phonemes aloud, proved especially successful, with all 24 items performing very well with a wide range of difficulty and high discrimination indices.

## 5. Discussion

Our first goal was to make sure that the test was usable (a test characteristic called *practicality* by Bachman & Palmer, 1996, p.35-37) and that young learners could indeed take the test autonomously on tablets. Teachers did not report major problems with test administration, scheduling, etc., and small groups of students were able to complete the test while classroom teachers continued teaching other groups. We conclude that the features we used to make the test more child-friendly (simple instructions, friendly monsters to contextualize the tasks) were successful, and that the test is usable by primary school teachers during class hours. Our results are in accordance with those of Carson et al. (2014, for English), Kiss & Csapó (2024, for Hungarian) and Meira et al. (2023, for Portuguese), who also developed computerized tests and used recorded, child-friendly instructions (Kiss & Csapó, 2024; Carson et al., 2014), child-appropriate pictures (Kiss & Csapó, 2024; Meira et al., 2023) or characters (an alien in Carson et al., 2014) to contextualize PA tasks and motivate the children. This confirms the usefulness of gamification features that are known to promote motivation, such as stories and characters or avatars (Sailer et al., 2017), when developing a test aimed at young children.

The second goal of this study was to make sure that our test was reliable, i.e. that the use of the test on separate occasions would yield similar results. As a proxy, we used a commonly accepted measure of internal consistency, Cronbach's alpha. Each of the subparts (tasks) composing our test displayed adequate levels of internal consistency. We can thus conclude that the reliability of the test is satisfactory. This confirms results of previous studies such as Kiss & Csapó (2024), who showed that a PA test taken autonomously on tablets could be reliable. However, the procedure for test administration was different in their case, since children took the test in groups of five in a quiet room with two adult assessors present to answer their questions and solve any problems. Our study shows that even in less controlled conditions with less supervision (children taking the test in small groups in the classroom while the teacher is working with other students) a computerized PA test can still exhibit satisfactory reliability. However, the students knew they were taking part in an experiment and this might have skewed the results. Replication with a more typical group would be useful.

Finally, our test was made up of different tasks (blending, segmentation, similarity and oddity) aimed at different phonological units (syllables, rhymes, units). Although tasks focusing on smaller units (phonemes) are usually found to be harder, we did not find evidence of such a trend in our study. For example, the two tasks aimed at syllables (syllable blending and syllable oddity) were respectively the easiest and the hardest tasks, and the phoneme segmentation task was easier than both the syllable oddity and the rhyme similarity tasks. Our results are in line with those of Kiss & Csapó (2024), who also found that task difficulty seems to be determined less by phonological unit size than by the cognitive operation required.

## 6. Conclusion, limitations and further research

This project has enabled us to develop a usable and reliable phonological awareness test that can be taken autonomously on a tablet by young learners. All four PA sub-tests proved to be reliable, of adequate difficulty, and discriminating well between low-performing and high-performing learners. However, students and their teachers taking part in this study were already familiar with the use of tablets. With students who are new to them, a period of familiarization might be necessary before taking the test itself.

The initial aim was to meet our research needs in applied linguistics, but also the needs of teachers to assess the level of their learners. The experiment highlighted the fact that teachers can administer the test on their own, and that this saves resources compared with a conventional test. As far as future developments for teachers are concerned, we now need to work on designing a dashboard on which teachers can consult the results obtained by learners. We have been able to provide evidence for several aspects of validity (Hughes, 2002). Apart from reliability we have also explored content validity, by showing that the test covers different aspects of phonological awareness, both in terms of task types, and in terms of the size of phonological units concerned. However, test validation is an ongoing process and there is still much work to be done to confirm these preliminary results.

In terms of research, we intend to conduct an external validation study to verify criterion reliability. To this end, we plan to administer our test as well as a traditional test recognized in the field by language specialists to a different group of students from the ones participating in this study. This will enable us to determine whether the results obtained in the two tests are correlated, i.e. whether the students who obtain the best results in our test are also those who obtain the best results in the traditional test. We will also be able to compare the time needed in both modes of administration, which was not possible in this study.

Furthermore, considering our research hypothesis that phonological awareness is not language-specific, we still need to compare the results on the French items and the English items (preliminary analyses not presented in this article suggest that they are highly correlated). We also plan to add items in Taiwanese. We will then be able to see whether the level of performance is the same whatever the language of the items.

## Acknowledgements

The authors acknowledge the support of ANR – FRANCE (French National Research Agency) for the C-maphore project n°21-CE38-0005 and the Trans3 project (France 2030, e-FRAN-vague3), ANR-22-FRAN-0008. They also thank the LABEX ASLAN of Université de Lyon (ANR-10-LABX-0081) for its financial support as part of the ‘Investissements d’Avenir’ program. We would also like to thank the editor for her insightful remarks.

## References

- ALTE (Ed.). (1998). *Multilingual glossary of language testing terms*. Cambridge University Press.
- Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, 14(5), 255–259. <https://doi.org/10.1111/j.0963-7214.2005.00376.x>
- Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology*, 96(1), 43–55. <https://doi.org/10.1037/0022-0663.96.1.43>
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge University Press.
- Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, 39(5), 913–923. <https://doi.org/10.1037/0012-1649.39.5.913>



- Carson, K., Boustead, T., & Gillon, G. (2014). Predicting reading outcomes in the classroom using a computer-based phonological awareness screening and monitoring assessment (Com-PASMA). *International Journal of Speech-Language Pathology*, 16(6), 552–561. <https://doi.org/10.3109/17549507.2013.855261>
- Charles, É., Magnat, É., Jouannaud, M.-P., Payre-Ficout, C., & Loiseau, M. (2022, September). *Effect of an EFL listening comprehension learning game on phonemic awareness in French*. ATFLY 2022 — Advances in Teaching Foreign Languages to Young Learners, online. <https://hal.archives-ouvertes.fr/hal-03865677>
- Comeau, L., Cormier, P., Grandmaison, É., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, 91(1), 29–43. <https://doi.org/10.1037/0022-0663.91.1.29>
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251. <https://doi.org/10.3102/00346543049002222>
- Gillon, G. T. (2018). *Phonological awareness: From research to practice* (Second edition). The Guilford Press.
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read* (pp. viii, 166). Lawrence Erlbaum Associates, Inc.
- Haigh, C. A., Savage, R., Erdos, C., & Genesee, F. (2011). The role of phoneme and onset-rime awareness in second language reading acquisition. *Journal of Research in Reading*, 34(1), 94–113. <https://doi.org/10.1111/j.1467-9817.2010.01475.x>
- Hoiem, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I.-K. (1995). Components of phonological awareness. *Reading and Writing*, 7(2), 171–188. <https://doi.org/10.1007/BF01027184>
- Hopp, H., Steinlen, A., Schelleter, C., & Piske, T. (2019). Syntactic development in early foreign language learning: Effects of L1 transfer, input, and individual factors. *Applied Psycholinguistics*, 40(5), 1241–1267. <https://doi.org/10.1017/S0142716419000249>
- Hughes, A. (2002). *Testing for Language Teachers*. Cambridge University Press.
- Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S., & Zorman, M. (2010). *Batterie Analytique du Langage Écrit*. Cogni-Sciences. <https://www1.ac-grenoble.fr/article/cognisciences-121593>
- Kersten, K., Jedamski, L., Bruhn, A.-C., & Bredel, U. (2024, July). *Testing a phonological awareness test: A valid reduced scale predicts L1 and L2 receptive lexical and grammatical skills*. EUROSLA 2024, Montpellier.
- Kiss, R., & Csapó, B. (2024). Technology-based assessment of phonological awareness in kindergarten. *International Journal of Early Childhood*. <https://doi.org/10.1007/s13158-023-00386-7>
- Kivistö de Souza, H. (2015). Phonological awareness and pronunciation in a second language [Ph.D. Thesis, Universitat de Barcelona]. In *TDX (Tesis Doctorals en Xarxa)*. <http://www.tdx.cat/handle/10803/393726>
- Laveault, D., & Grégoire, J. (2014). *Introduction aux théories des tests en psychologie et en sciences de l'éducation*. De Boeck.
- Li, M., Cheng, L., & Kirby, J. (2013). Phonological awareness and listening comprehension among Chinese English-immersion students. *International Education*, 41(2). <https://trace.tennessee.edu/internationaleducation/vol41/iss2/4>
- Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201–212. [https://doi.org/10.1016/0022-0965\(74\)90101-5](https://doi.org/10.1016/0022-0965(74)90101-5)
- Magnat, E. (2013). *Le TBI comme instrument du développement de la conscience phonémique à l'école: Une approche ergonomique* [These de doctorat, Grenoble]. <http://www.theses.fr/2013GRENL005>
- Mandin, S., Zaher, A., Meyer, S., Loiseau, M., Bailly, G., Payre-Ficout, C., Diard, J., Valdois, S., Blavot, A., Bosse, M.-L., Briswalter, Y., Chalon, N., Godde, E., Ingremeau, S., Jouannaud, M.-P., Lequette, C., Magnat, E., Masperi, M., Piat-Marchand, A.-L., ... Zanoni, M. (2021). Expérimentation à grande échelle d'applications pour tablettes pour favoriser l'apprentissage de la lecture et de l'anglais oral. In M. Lefèvre, C. Michel, T. Geoffre, M. Rodi, L. Alvarez, & A. Karoui (Eds.), *Actes de la 10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain* (pp. 118–129). ATIEF. <https://hal.archives-ouvertes.fr/hal-03292798>

- Meira, Â., Cadime, I., & Viana, F. L. (2023). Phonological Awareness Assessment Test (PACOF) for pre-school children: Evidence of validity and reliability. *Educational Psychology*, 29(1), 83–90. <https://doi.org/10.5093/psed2023a3>
- Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, 34(1), 114–135. <https://doi.org/10.1111/j.1467-9817.2010.01477.x>
- Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, 69, 371–380. <https://doi.org/10.1016/j.chb.2016.12.033>
- Sprenger-Charolles, L., Colé, P., Béchenec, D., & Kipffer-Piquard, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests. *European Review of Applied Psychology*, 55(3), 157–186. <https://doi.org/10.1016/j.erap.2004.11.002>
- White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness*, 11(4), 259–290. <https://doi.org/10.1080/09658410208667060>
- Wills, S., Bai, Y., Tejedor-Garcia, C., Cucchiarini, C., & Strik, H. (2023). Automatic speech recognition of non-native child speech for language learning applications. *OASICs, Volume 113, SLATE 2023*, 113, 7:1-7:8. <https://doi.org/10.4230/OASICs.SLATE.2023.7>