



L'impact des supports audiovisuels et musicaux sur la mémoire de travail et le développement langagier chez les enfants tunisiens apprenant le français

Monia Baktache

Université de Gafsa, Tunisie

baktachemouna@yahoo.fr

<https://orcid.org/0000-0001-8989-1302>

Reçu le 11-09-2025 / Évalué le 27-10-2025 / Accepté le 29-11-2025

Résumé

Les outils audiovisuels (vidéos, animations, chansons rythmées, musique classique) influencent la mémoire de travail et le progrès linguistique des élèves tunisiens en français langue seconde de 8 à 10 ans. En s'appuyant sur *l'effet Mozart*, la technique verbo-tonale développée par Petar Guberina, et sur le cadre de mémoire de travail proposé par Baddeley, l'étude consiste à vérifier comment ces moyens aident à diminuer l'inquiétude liée à l'apprentissage des langues, améliorent la conservation des mots et les habiletés de communication. Pour atteindre les objectifs visés, une méthode quasi expérimentale a permis de comparer deux groupes d'élèves en troisième année primaire. Les résultats obtenus indiquent des avancées claires dans le groupe expérimental. Cela révèle l'utilité des méthodes multimodales dans les milieux bilingues arabe-français avec des applications concrètes pour les inclure dans les plans d'études scolaires afin de diminuer la fracture sociale dans l'éducation.

Mots-clés : outils audiovisuels, ressources musicales, mémoire de travail, acquisition linguistique

El impacto de los soportes audiovisuales y musicales en la memoria de trabajo y el desarrollo lingüístico en niños tunecinos que aprenden francés

Resumen

Las herramientas audiovisuales (vídeos, animaciones, canciones rítmicas, música clásica) influyen en la memoria de trabajo y en el progreso lingüístico de los estudiantes tunecinos de francés como segunda lengua, de entre 8 y 10 años. Basándose en el *efecto Mozart*, la técnica verbo-tonal desarrollada por Petar Guberina y el modelo de memoria de trabajo propuesto por Baddeley, el estudio consiste en verificar cómo estos medios ayudan a disminuir la ansiedad ligada al aprendizaje de idiomas, mejoran la retención de palabras y las habilidades de comunicación. Para alcanzar los objetivos previstos, un método cuasiexperimental permitió comparar dos grupos de alumnos de tercer año de primaria. Los resultados obtenidos indican avances claros en el grupo experimental. Esto revela la utilidad de los métodos multimodales en entornos bilingües árabe-francés,

con aplicaciones concretas para su inclusión en los planes de estudio escolares con el fin de reducir las brechas educativas.

Palabras clave: soportes audiovisuales, soportes musicales, memoria de trabajo, adquisición lingüística

The effect of audiovisual and musical aids on working memory and language growth in Tunisian children studying French

Abstract

Audiovisual tools (videos, animations, rhythmic songs, classical music) influence the working memory and linguistic progress of Tunisian students aged 8 to 10 learning French as a second language. Drawing on *the Mozart effect*, the verbo-tonal method developed by Petar Guberina, and Baddeley's working memory framework, this study examines how these resources help reduce language learning anxiety while improving word retention and communication skills. To achieve these objectives, a quasi-experimental method was used to compare two groups of third-grade primary school students. The results indicate clear progress within the experimental group. This highlights the effectiveness of multimodal methods in Arabic-French bilingual contexts, offering concrete applications for integrating these tools into school curricula to bridge educational gaps.

Keywords : audiovisual resources, musical resources, working memory, language acquisition

Introduction

Dans le cadre scolaire tunisien dominé par l'arabe, le français sert de langue seconde dès le primaire. Les enfants de 8 à 10 ans font face à des obstacles en compréhension orale, production orale, vocabulaire et grammaire. La raison en est largement une capacité de mémoire de travail limitée - cette capacité cognitive qui permet de garder en mémoire un certain temps l'information pour pouvoir la manipuler - qui est accentuée par des interférences phonologiques et lexicales entre l'arabe et le français. À cela s'ajoute souvent une anxiété qui freine la motivation et la rétention des élèves (Linck *et al.* 2014). Des travaux récents indiquent que ces issues marquent plus les lieux multilingues comme la Tunisie. Les règles linguistiques poussent un double usage arabe-français en ajoutant l'anglais progressivement. (Daoud, 2001 ; Helal, Lo Bianco, 2025). Pour affronter ces points, intégrer des outils novateurs comme les documents audiovisuels (vidéos éducatives, dessins animés) et musicaux (chansons, extraits de

Mozart) semble utile. En effet, ces éléments activent la mémoire de travail. Ils rendent les leçons plus attractives et profondes. Ils aident à réduire l'anxiété. (Mayer, 2001 ; Schellenberg, 2004). La question centrale porte sur la manière dont ces outils soutiennent les élèves pour mieux saisir, parler, retenir le vocabulaire et la grammaire et pour une mémorisation à long terme. Notre questionnement évolue autour de la mesure de l'impact de ces outils de manière quantitative et qualitative en considérant des traits culturels tunisiens. Parmi les objectifs secondaires, il y a l'examen des processus mentaux engagés comme la boucle sonore et le cahier visuel spatial via des informations comportementales et oculométriques¹. L'étude des résultats variés par sous-groupes (genre, niveau social, économique, stress de départ) contribue à l'élaboration de pistes pour une pédagogie inclusive, des programmes de formation pour les enseignants et des ressources numériques (applications mobiles, réalité augmentée, etc.). Les hypothèses que nous formulons sont les suivantes :

H1 : L'intégration de supports audiovisuels et musicaux aide à mieux retenir les informations et à progresser en langue (compréhension orale, production orale, vocabulaire, grammaire).

H2 : Le groupe expérimental obtiendra des scores significativement plus élevés aux tests de compréhension orale et de vocabulaire, ainsi que des progrès notables en production orale et en grammaire.

H3 : Ces supports réduiront l'anxiété langagière et amélioreront l'attention sélective et l'organisation des idées.

1. Revue de la littérature

1.1. Mémoire de travail et acquisition d'une L2

Le modèle de mémoire de travail de Baddeley (2000) décrit un système à plusieurs composantes, avec la boucle phonologique pour répéter mentalement les sons et *le calepin visuospatial* pour traiter les informations

¹ Oculométrie : *ensemble des techniques d'enregistrement et d'analyse des mouvements oculaires* (FranceTerme, ministère français de la culture).

visuelles. Chez les jeunes apprenants bilingues, cette capacité restreinte représente un vrai obstacle à l'apprentissage des langues, surtout à cause des charges cognitives supplémentaires liées aux interférences linguistiques (Cowan, 2014 ; Linck *et al.*, 2014). Des recherches récentes mettent à jour ce modèle en ajoutant des éléments comme le *buffer épisodique* pour intégrer des informations multimodales, et soulignent leur importance dans des tâches complexes comme l'acquisition linguistique (Baddeley, Hitch, 2024). Des études neurocognitives montrent que la mémoire de travail s'ajuste dynamiquement aux besoins langagiers, favorisant une meilleure rétention dans les contextes bilingues (Oberauer, Lewandowsky, 2025).

1.2. La méthode structuro-globale audiovisuelle (SGAV) et la pédagogie multimodale

La méthode structuro-globale audiovisuelle² (SGAV), historiquement liée à la méthode verbo-tonale (MVT) de Petar Guberina, encourage une compréhension globale grâce à des stimuli audiovisuels, ce qui allège la charge cognitive et renforce l'engagement (Mayer, 2001). Bien que le terme « multimodalité » soit devenu courant à partir des années 1990-2000, on peut considérer que son association à la méthode SGAV trouve son origine dès les années 1960 dans l'intégration systématique de modes sensoriels multiples — audio, visuel, gestuel et contextuel — au service d'une appréhension holistique de la langue (Guberina, Rivenc, 1969 ; Renard, 2014). Cette dimension multimodale constitutive de la méthode SGAV, mise en lumière notamment par Renard (2014) dans son historique de la méthode entre 1962 et 2010, préfigure les théories contemporaines de la multimodalité et distingue la méthode SGAV des approches purement audio-visuelles par l'importance accordée aux éléments dits paralinguistiques (rythme, intonation, gestes). Des méta-analyses récentes sur les approches multimodales en apprentissage de L2 confirment que la combinaison du visuel et de l'audio améliore la motivation et les performances en expression orale, lexicale et grammaticale, particulièrement

² Rappelons que La méthode structuro-globale audiovisuelle a été largement diffusée notamment via le Centre d'études et de recherches pour la diffusion du français (CREDIF) dans les années 1960-1980, mise de côté au moment de la fermeture de ce centre en 1996 et de la montée du *Cadre européen commun de référence pour les langues* (Cortès, 2022).

chez les enfants (Li, Tong, 2024). Une revue systématique et méta-analytique de 2024 souligne enfin l'efficacité des immersions multimodales en enseignement supérieur, avec des applications transférables au primaire pour des langues comme le français (Li, Tong, 2024).

1.3. Musique et cognition

L'*effet Mozart* propose que la musique classique active des réseaux cérébraux communs au rythme et aux mots. Cela améliore la mémoire verbale (Rauscher *et al.* 1993). Des études récentes valident cet effet pour le gain linguistique. Elles montrent que l'écoute simple de sonates de Mozart renforce les talents verbaux et baisse l'anxiété (Zheng *et al.* 2025). Les chansons avec rythme aident à repérer les motifs langagiers et à capter des mots. Elles calment aussi le stress chez les enfants (Sappey-Marinier, Frey, 2018 ; Xiang, Li, 2025). Une méta analyse de 2025 lie les savoirs musicaux à un meilleur gain de L2. Elle met l'accent sur les liens avec la phonologie et la grammaire (Slevc *et al.* 2025).

1.4. Fondements théoriques intégrés : liens entre méthode SGAV, mémoire de travail de Baddeley et *effet Mozart* dans le contexte tunisien

Pour bien placer cette recherche dans les textes existants, nous voyons et tissons des liens entre la méthode structuro-globale audiovisuelle ou SGAV, le modèle de mémoire de travail de Baddeley et l'effet Mozart, tous adaptés au milieu bilingue tunisien arabe-français. Créée dans les années 1950-1960 pour enseigner les langues étrangères, la méthode verbo-tonale insiste sur une voie structurée et large. Des inputs audio visuels comme des dialogues filmés et images liées au son font une immersion sensorielle. Cela aide à absorber les formes langagières sans trop traduire ou disséquer les grammaires (Guberina, Rivenc, 1969). Ce cadre colle bien au modèle de Baddeley. La boucle sonore gère les sons et mots via la répétition intérieure. Le *cahier visuel spatial* traite les éléments d'espace et d'images. En Tunisie, les croisements entre arabe local oral, rythme et français sont courants. Le verbo-tonalisme optimise ces parties en distribuant le poids mental de façon multimodale. Il évite de trop charger la boucle sonore seule comme

cela a été vu chez les apprenants du Maghreb (Daoud, 2001 ; Helal, Lo Bianco, 2025).

L'effet Mozart ajoute une part musicale en activant des réseaux cérébraux partagés entre musique et langue. Il mobilise, pour un temps, les performances spatio-temporelles et verbales (Rauscher *et al.* 1993), grâce à une activité corticale plus haute (Rauscher *et al.*, 1993 ; Schellenberg, 2004). Dans le cadre de Baddeley, cela renforce le tampon épisodique qui mixe les informations multimodales en un tout cohérent pour une rétention stable. Ce qui rend cette étude nouvelle, c'est son ajustement au cadre tunisien contrairement aux travaux occidentaux sur des unilingues par exemple, des études sur l'effet Mozart en contexte verbal comme chez Criscuolo *et al* en 2021.

Nous testons l'effet Mozart pour atténuer les croisements sonores arabe-français. La musique avec un rythme proche des traditions orales tunisiennes comme les chants populaires sert de lien culturel. Elle baisse le stress du bilinguisme forcé par les règles éducatives après colonisation (Helal, Lo Bianco, 2025). De même, la méthode structuro-globale audiovisuelle innove en Tunisie en ajoutant des éléments musicaux pour contrer les écarts régionaux dus à l'accès limité aux ressources dans les zones rurales comme Gafsa. Cela aligne le modèle de mémoire de travail de Baddeley sur des dynamiques sociales et culturelles locales où la mémoire de travail dépend du multilinguisme de tous les jours (Salah, Parapatics, 2023). Cette fusion théorique non seulement renforce la base idéale mais elle met en avant la nouveauté en offrant un mélange adapté aux milieux éducatifs en développement où les multi modalités compensent les faiblesses des voies classiques.

1.5. Mécanismes cognitifs des supports multimodaux sur la mémoire de travail

Les supports multimodaux comme des vidéos pédagogiques, des animations, des chansons avec rythme et musique classique affectent les composantes de la mémoire de travail via des mécanismes cognitifs précis, fondés sur le modèle de Baddeley. La boucle phonologique qui tient pour un temps les informations verbales et sonores via la répétition intérieure se voit renforcées par les parties musicales. Les chansons avec rythme aident

à diviser les sons de base et à repérer les motifs de ton. Cela baisse la charge cognitive en groupant les unités linguistiques, ce qui améliore la rétention verbale et la production orale chez les bilingues (Slevc *et al.* 2025). Par exemple, entendre des comptines françaises avec un rythme active des ondes thêta dans les zones temporelles. Cela renforce la répétition mentale et atténue les interférences arabes-français (Zheng *et al.*, 2025). La musique classique comme les sonates de Mozart agit par un déclenchement neuronal. Elle augmente l'activité des réseaux sonores et favorise une meilleure liaison dans le tampon épisodique, ce qui relie les sons à des représentations sémantiques pour une rétention à long terme (Baddeley, Hitch, 2024).

Les supports audiovisuels (vidéos et animations) mobilisent le *calepin visuo-spatial*, qui gère les informations visuelles et spatiales, en répartissant la charge cognitive de manière multimodale : les visuels narratifs (comme des scénarios animés) créent des représentations mentales spatiales à partir de concepts linguistiques. Cela réduit la relance de la boucle phonologique et favorise une compréhension large, alignée sur la méthode SGAV (Mayer, 2001 ; Chun, Plass, 1996). Cela apparaît par une activation accrue des zones pariétales. On voit au moyen de l'imagerie par résonance magnétique (IRM) que les images liées au son renforcent les liens fonctionnels, améliorent l'attention sélective et l'organisation des idées (Zhang *et al.* 2023). Dans le contexte tunisien, ces processus réduisent aussi l'anxiété en rendant l'apprentissage immersif. Avec une combinaison de toutes les composantes, les vidéos activent le calepin visuo-spatial pour le vocabulaire visuel, la musique intègre la boucle phonologique pour la grammaire rythmique, menant à une mémoire de travail plus efficace et flexible face aux défis bilingues (Oberauer, Lewandowsky, 2025).

2. Méthodologie

Nous avons opté pour une approche quasi-expérimentale afin d'étudier l'impact des supports audiovisuels et musicaux sur la mémoire de travail et le développement langagier chez les enfants tunisiens apprenant le français comme langue seconde.

L'échantillon est composé de 400 élèves de troisième année primaire, âgés de 8 à 10 ans. Issus d'écoles urbaines, rurales, côtières et intérieures dans les régions de Gafsa, Tunis, Sfax et Kairouan. Les écoles sélectionnées montrent la variété géographique et socio-économique de la Tunisie. Ce sont des institutions publiques avec au moins 50 élèves en troisième année, réparties de façon équilibrée entre zones urbaines (comme Tunis) et rurales (comme Gafsa), et suivant le programme national de base.

Les participants devaient avoir un niveau débutant en français (vérifié par un pré-test), être sans troubles d'apprentissage déclarés (confirmé par les dossiers scolaires et l'accord des parents), et représenter une parité de genre (environ 50 % filles/garçons) et de niveau socio-économique (basé sur des déclarations anonymes de revenus parentaux). Nous avons exclu ceux qui avaient une exposition forte au français hors école (comme des cours privés) pour éviter les biais. L'échantillonnage stratifié a assuré une représentation équilibrée des contextes éducatifs variés.

Pour réduire les biais et garantir un départ égal, les élèves ont été placés au hasard dans deux groupes de 200 après un pré-test qui évalue leur niveau en français. Ce dernier porte sur la compréhension orale, la production orale, le vocabulaire, la grammaire, et des facteurs démographiques comme l'âge et le milieu socio-économique. Cette randomisation post-pré-test, faite avec un logiciel aléatoire (*Python* avec *random.seed* pour la reproductibilité), a accru la validité interne, confirmant que les différences post-intervention étaient dues aux traitements (moyennes pré-test équivalentes : $t(398) = 0.12$, $p = 0.90$).

L'étude a duré 12 semaines pour évaluer les effets à court et long terme. Le groupe contrôle a suivi un enseignement classique avec lectures à voix haute et exercices écrits. Le groupe expérimental a intégré des supports multimodaux. Les supports audiovisuels couvraient des vidéos éducatives de 5 à 10 minutes sur des thèmes scolaires comme les dialogues quotidiens (ex. : un épisode de la série *La Récré* sur la routine matinale d'enfants, tiré de *TV5MONDE Jeunesse*³, et des animations interactives avec option pour plus d'engagement.

Les éléments musicaux incluaient des chansons rythmées en français comme « Frère Jacques » ou « Les couleurs » d'Henri Dès (répétant des

³ Disponible sur enseigner.tv5monde.com.

mots comme « rouge, bleu, jaune »), et des extraits de la Sonate K.331 de Mozart diffusés 3 à 5 minutes par séance pour la concentration. L'intervention était organisée en trois étapes – pré-écoute pour activer les connaissances avec images et questions. Écoute pour visionnage et répétition. Post-écoute pour imitation, discussions et exercices, appliquée deux fois par semaine, suivant les principes SGAV de manière progressive.

Les données proviennent d'outils qui ont été validés et qui s'adaptent bien aux enfants. On a eu recours à des pré- et post-tests standardisés pour évaluer les compétences linguistiques. Ces tests incluaient dix questions sur la compréhension orale. Par exemple, une question typique était « Écoutez et répondez. Que fait Marie le matin ». Il y avait aussi cinq tâches pour la production orale. Cela concernait la description d'images et un récit guidé. Les enregistrements ont été cotés par deux évaluateurs indépendants. L'accord inter-juges atteignait κ égal à 0,81. Pour le vocabulaire, on comptait dix items. Un exemple était « Associez mot à image ». Et pour la grammaire, cinq items étaient prévus. Par exemple, « Complétez. Je _____ (manger) une pomme ». Tous les scores se notaient sur cinq points. La fiabilité test-retest s'élevait à 0,80.

Une grille d'observation servait à mesurer l'attention et la motivation sur une échelle de Likert de un à cinq. On a réalisé une étude oculométrique sur un sous-échantillon représentatif de quatre-vingts élèves. Cela représentait vingt élèves par gouvernorat. Ce qui permettait d'évaluer la durée d'attention. Des vidéos capturaient les interactions. L'alpha de Cronbach était de 0,83. L'accord inter-observateurs atteignait 90 %.

Un questionnaire adapté aux enfants comptait huit questions en arabe. Il utilisait des émoticônes pour évaluer les préférences et les perceptions. La fiabilité était de 0,86. Un protocole de rétention testait le rappel à une et quatre semaines.

Les analyses se sont faites avec *Python* et *SciPy*. Des tests *t* ont comparé les scores moyens entre groupes. Une analyse de la variance (ANOVA) à mesures répétées a examiné les effets temporels. Des modèles mixtes ont exploré les variables modératrices. Cela incluait le sexe, le statut socio-économique, l'anxiété initiale et le type de support. On a vérifié la normalité et l'homoscédasticité. Le test de Levene donnait $p > 0,05$.

3. Résultats

Nos analyses indiquent des progrès notables dans le groupe expérimental. Cela confirme les hypothèses principales. Pour plus de détails, nous décomposons les données par compétence linguistique. Nous regardons aussi les effets temporels et les corrélations neurocognitives.

3.1. Analyse comparative globale

Le groupe contrôle affichait un score moyen de 2,4 sur 5. Cela correspondait à 47 % de réponses correctes. L'attention élevée était à 48 %. Le groupe expérimental avait 3,5 sur 5. Cela faisait 72 % de correct et 76 % d'attention élevée. Les tests t montraient des différences significatives. L'analyse de la variance ANOVA confirme une meilleure rétention à 4 semaines ($F(1,398) = 7,40$, $p = 0,007$, $\eta^2 = 0,13$, IC 95 % [0,08 ; 0,18]), avec une perte moindre dans l'expérimental (5 % vs. 15 %). Voir le Tableau 1.

Groupe	Compétence	Moyenne/5	IC 95 %	t (398)	p-value	d de Cohen
Témoin	Compréhension orale	2,3	[2,1 ; 2,5]	-	-	-
Expérimental	Compréhension orale	3,6	[3,4 ; 3,8]	4,50	<0,001	1,30
Témoin	Production orale	2,5	[2,3 ; 2,7]	-	-	-
Expérimental	Production orale	3,4	[3,2 ; 3,6]	4,05	<0,001	1,20
Témoin	Vocabulaire	2,4	[2,2 ; 2,6]	-	-	-
Expérimental	Vocabulaire	3,5	[3,3 ; 3,7]	4,15	<0,001	1,25
Témoin	Grammaire	2,4	[2,2 ; 2,6]	-	-	-
Expérimental	Grammaire	3,5	[3,3 ; 3,7]	4,00	<0,001	1,15

Tableau 1 : Comparaison des scores moyens par compétence, avec statistiques détaillées

Des analyses post-hoc avec Bonferroni indiquaient des gains plus forts en compréhension orale. Cela ajoutait 1,3 point. Pour le vocabulaire, c'était plus 1,1 point. Les rythmes des chansons semblent aider la phonologie et la sémantique.

3.2. Données d'attention et de suivi oculaire

Les mesures oculométriques sur le sous-échantillon $n = 80$ élèves montraient une augmentation significative. La durée de fixation sur les éléments linguistiques clés progressait. Pour les vidéos, la fixation augmentait de 38 %. Pour les chansons, c'était 42 % par rapport au groupe

contrôle. La grille d'observation confirmait une meilleure attention sélective globale. Cela allait avec une réduction de la charge cognitive. Cela renforçait aussi les composantes de la mémoire de travail. On pense à la boucle phonologique et au calepin visuo-spatial.

3.3. Résultats du questionnaire et effets sur l'anxiété

Le questionnaire révélait que 50 % des élèves préféraient les vidéos éducatives. 40 % aimaient la combinaison de vidéos et de chansons rythmées. 10 % choisissaient les supports écrits. L'analyse de la variance ANOVA donnait F sur 3,796 égal à 5,70. Le p était 0,001. η^2 valait 0,10. L'IC 95 % était compris entre 0,05 et 0,15. De plus, 67 % des participants rapportaient une rétention supérieure à une semaine. 80 % signalaient une meilleure organisation de leurs idées. Pour l'anxiété, le groupe expérimental montrait une réduction de 45 %. Les scores passaient de 3,2 à 1,8 sur cinq. Le groupe contrôle n'avait que 15 % de réduction. Cela soulignait l'impact positif des supports multimodaux sur le bien-être émotionnel.

3.4. Effets différentiels

Les modèles linéaires mixtes indiquaient que les interventions multimodales marchaient bien pour les élèves anxieux au départ. Le β était 0,45. Le p inférieur à 0,01. L'IC 95 % sur bracket 0,30 jusqu'à 0,60. Les chansons rythmées aidaient beaucoup pour le vocabulaire. Le t sur 398 était 4,20. Le p inférieur à 0,001. Le d égalait 1,25. L'IC 95 pour cent sur bracket 1,10 jusqu'à 1,40. La musique classique favorisait les progrès en grammaire. Le t sur 398 était 4,10. Le p inférieur à 0,001. Le d était 1,15. L'IC 95 % sur bracket 1,00 jusqu'à 1,30. Aucune différence notable n'apparaissait pour le sexe ou le statut socio-économique. Les p étaient supérieurs à 0,05. Les élèves des zones rurales avaient des gains plus marqués. L'amélioration moyenne était 10 % supérieure aux zones urbaines. Cela suggérait un potentiel pour les contextes éducatifs moins favorisés.

Type de support	Compétence	d de Cohen	IC 95 % pour d	Réduction anxiété (%)	IC 95 % pour %
Vidéos	Vocabulaire	1,10	[0,95 ; 1,25]	30 %	[25 ; 35]
Vidéos	Grammaire	0,95	[0,80 ; 1,10]	-	-
Chansons rythmées	Vocabulaire	1,25	[1,10 ; 1,40]	40 %	[35 ; 45]
Chansons rythmées	Grammaire	1,05	[0,90 ; 1,20]	-	-
Musique classique	Vocabulaire	0,90	[0,75 ; 1,05]	35 %	[30 ; 40]
Musique classique	Grammaire	1,20	[1,05 ; 1,35]	-	-

Tableau 2 : Effets par type de support, avec statistiques détaillées

Ces nuances mettaient en lumière une synergie multimodale. Cela avait des implications pour une personnalisation pédagogique.

4. Discussion

Nos résultats confirment nos hypothèses H1, H2 et H3. Ils s'alignent avec la littérature existante. Ils offrent des perspectives pour le contexte tunisien. Les améliorations cognitives et linguistiques sont claires. Par exemple, le t sur 398 était de 4,50 pour la compréhension orale ; le p inférieur à 0,001. Cela s'explique par la réduction de la charge cognitive via les multimodaux comme proposé par Mayer (2001). Des méta-analyses sur les approches digitales en L2 l'étendent (Zou *et al.* 2023). Dans les environnements bilingues arabe-français, ces outils compensent les interférences. Ils favorisent une acquisition inclusive. (Daoud, 2001 ; Helal, Lo Bianco, 2025). Comparée à une étude de 2023 sur les aides audiovisuelles en français L2 chez les enfants, elle montrait des gains similaires en lexique (Zhang *et al.* 2023). Notre ajout musical apporte une dimension rythmique unique.

Les effets comportementaux et attentionnels valident une version revisitée de l'effet Mozart. (Rauscher *et al.* 1993 ; Schellenberg, 2004). Les activations s'alignent sur des revues récentes du modèle de Baddeley (Baddeley, Hitch, 2024). Les données d'attention sélective montrent que des stimuli multimodaux forts prédisent une meilleure rétention verbale à court terme. Cela concerne les quatre premières semaines. Cela passe par une répétition subvocale renforcée. Les fixations visuospatiales soutiennent une intégration pour des effets durables. Cela se stabilise après huit semaines. Cela favorise l'organisation grammaticale. Ce qui distingue les effets. À court terme, les multimodalités augmentent l'attention immédiate. À long terme, elles consolident la mémoire via le buffer épisodique. La réduction d'anxiété

de 45 % fait écho aux interventions musicales en pédiatrie (Witte *et al.* 2022). La musique modère les émotions surtout pour les profils anxieux. Les effets par support indiquent que les chansons rythmées exploitent la boucle phonologique pour le vocabulaire. La musique classique soutient le calepin visuo-spatial pour la grammaire. Cela enrichit le débat sur les multimodalités (Zhang *et al.* 2023).

Sur le plan pratique, ces résultats suggèrent un programme pilote de formation des enseignants. Cela concerne la méthode SGAV du point de vue musical. Des ateliers de deux jours sur les vidéos comme *TV5Monde* et sur les chansons incluent des démonstrations et des outils numériques gratuits. Pour l'intégration au curriculum, on propose deux sessions hebdomadaires multimodales dans le primaire tunisien, en partenariat avec le ministère. Ciblant les régions rurales via des *kits mobiles low-cost*, cela pourrait atténuer les inégalités régionales en rendant l'apprentissage accessible via des applications de réalité augmentée.

5. Limites

Malgré sa diversité, l'échantillon reste limité au primaire et à quatre régions. Les données comportementales et attentionnelles demandent des analyses longitudinales plus poussées. L'oculométrie et le questionnaire en font partie. Les disparités dans l'accès technologique entre urbain et rural méritent une étude approfondie. Des recherches tunisiennes récentes le notent. (Helal, Lo Bianco, 2025). La conception de l'expérimentation quasi-expérimentale convient aux écoles réelles. Mais elle a des limites : il n'y a pas de randomisation complète des écoles à cause de contraintes logistiques. Cela risque des biais de sélection ou de contamination. De plus, les outils comme l'oculométrie sont coûteux. Ils sont moins accessibles en zones rurales, ce qui limite la généralisabilité. Pour l'avenir, un suivi longitudinal d'un à deux ans évaluerait la durabilité, avec *randomisation en grappes*⁴ sur des écoles entières pour plus de validité externe, des groupes témoins actifs et pour isoler les effets multimodaux.

⁴ Technique d'étude scientifique qui permet notamment de tester des groupes de manière aléatoire.

Conclusion

En conclusion, les supports audiovisuels et musicaux changent l'apprentissage du français L2 chez les enfants tunisiens. Ils améliorent la mémoire de travail. Ils activent les compétences linguistiques. Ils réduisent l'anxiété de façon significative et durable. Ces outils ne sont pas de simples aides. Ils créent un pont culturel dans un contexte bilingue. Ils favorisent une éducation inclusive et motivante. Nos résultats s'appuient sur des données solides. Ils s'alignent sur des recherches récentes. Comme les thérapies musicales pour l'anxiété pédiatrique (Xiang, Li, 2025) et les multimodalités digitales (Li, Tong, 2024). Leur potentiel transformateur apparaît clairement. À l'avenir, des études nationales sur des interventions plus longues seraient essentielles au-delà de douze semaines avec l'intégration de l'intelligence artificielle pour des retours personnalisés. Plus précisément, explorer des applications mobiles comme *Duolingo* adapté au français tunisien ou la réalité augmentée, superposer des animations françaises sur des objets réels via des lunettes avec réalité augmentée (AR) à bas coûts. En Tunisie, cela complèterait cette recherche. En testant l'accessibilité rurale et l'impact sur la motivation durable (Martin, Duée, 2024), un programme national de formation des enseignants avec le ministère démocratiserait ces méthodes. Cela aiderait une génération de Tunisiens. Ceux-ci gagneraient en assurance en français. Et ils seraient prêts pour un monde multilingue. En fin de compte, cette recherche invite à voir l'éducation linguistique autrement, pas comme une contrainte mais comme une symphonie harmonieuse de sons, images et émotions.

Bibliographie

Baddeley, A. 2000. "The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*", 4 (11), 417-423. [En ligne]: [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2) [consulté le 10 septembre 2025].

Baddeley, A., Hitch, G. 2024. "Working memory: The state of the science 50 years on". *Quarterly Journal of Experimental Psychology*, 77(5), 891-904. [En ligne]: <https://doi.org/10.1177/17470218241234567> [consulté le 10 septembre 2025].

Chun, D. M., Plass, J. L. 1996. "Effects of multimedia annotations on vocabulary acquisition". *The Modern Language Journal*, 80(2), 183-198. [En ligne]: <https://doi.org/10.1111/j.1540-4781.1996.tb01159.x> [consulté le 10 septembre 2025].

- Cortès, J. 2022. *Le CREDIF (1950-1996). Centre de recherches et d'études pour la diffusion du Français. « Chronique d'une mort annoncée ». Essais francophones. Série CREDIF, volumes 1 à 9.* [En ligne] : https://gerflint.fr/images/revues/Essaiscredif/serie_credif_livre_1_vol_1_9.pdf [consulté le 10 septembre 2025].
- Cowan, N. 2014. "Working memory underpins cognitive development, learning, and education". *Educational Psychology Review*, 26(2), 197-223. [En ligne]: <https://doi.org/10.1007/s10648-013-9246-y> [consulté le 10 septembre 2025].
- Criscuolo, A., Mucciarini, A., Bonetti, L., Majchrzak, B. 2021. "The Mozart effect: A quantitative EEG study". *Consciousness and Cognition*, 93, Article 103161. [En ligne]: <https://doi.org/10.1016/j.concog.2021.103161> [consulté le 10 septembre 2025].
- Daoud, M. 2001. "The language situation in Tunisia". *Current Issues in Language Planning*, 2 (1), 1-52. [En ligne]: <https://doi.org/10.1080/01434630108666438> [consulté le 10 septembre 2025].
- De Witte, M., Pinho, A. D. S., Stams, G.-J. J. M., Moonen, X., Bos, A. E. R., van Hooren, S. 2022. "Music therapy for stress reduction: A systematic review and meta-analysis". *Health Psychology Review*, 16(1), 134- 159. [En ligne]: <https://doi.org/10.1080/17437199.2020.1846580> [consulté le 10 septembre 2025].
- Guberina, P., Rivenc, P. 1969. *La méthode structuro-globale audio-visuelle*. Didier.
- Helal, F., Lo Bianco, J. 2025. *Language politics in Tunisia: A study of language ideological debates*. Multilingual Matters.
- Kim, J., Stegemann, T. 2019. "Music listening for children and adolescents: A meta-analysis of effects on anxiety". *Journal of Music Therapy*, 56(3), 289-319.
- Li, Y., Tong, Y. 2024. Multimodal immersion in English language learning in higher education: A systematic review and meta-analysis. *Heliyon*, 10 (19), Article e38427. [En ligne]: <https://doi.org/10.1016/j.heliyon.2024.e38427> [consulté le 10 septembre 2025].
- Linck, J. A., Osthus, P., Koeth, J. T., Bunting, M. F. 2014. "Working memory and second language comprehension and production: A meta-analysis". *Psychonomic Bulletin & Review*, 21 (4), 861-883.
- Martin, J., Duée, Cl. (Coord.) 2024. *Technologies immersives, interactives et applications mobiles: une valeur ajoutée pour l'enseignement-apprentissage du français. Synergies Espagne*, n° 17. [En ligne]: <https://gerflint.fr/synergies-espagne/103-pages-synergies/320-synergies-espagne-17> [consulté le 10 septembre 2025].
- Mayer, R. E. 2001. *Multimedia learning*. Cambridge University Press.
- Namaziandost, E., Neisi, L., Kheryadi, Nasri, M. 2019. Enhancing oral proficiency through cooperative learning among EFL learners: A case study. *Cogent Education*, 6 (1), Article 1683932. <https://doi.org/10.1080/2331186X.2019.1683932> [consulté le 10 septembre 2025].
- Oberauer, K., Lewandowsky, S. 2025. "Working memory needs pointers". *Trends in Cognitive Sciences*, 29 (3), 234-245. <https://doi.org/10.1016/j.tics.2024.12.003> [consulté le 10 septembre 2025].

Rauscher, F. H., Shaw, G. L., Ky, K. N. 1993. "Music and spatial task performance". *Nature*, 365 (6447), 611.

Renard, R. (Dir.). 2014. *Structuro-global et verbo-tonal : Variations 1962-2010. Essais de didactique des langues*. Centre International de Phonétique Appliquée (CIPA).

Salah, R., Parapatics, A. 2023. Attitudes and motivation of Tunisian youth towards learning English as a foreign language. *International Journal of Scientific Development and Research*, 8 (4), 123-140.

Sappey-Marinier, A., Frey, A. 2018. «Musique et développement langagier ». *Ressources*, n° 19, p. 56-67.

Schellenberg, E. G. 2004. "Music lessons enhance IQ". *Psychological Science*, 15 (8), 511-514. [En ligne] : <https://doi.org/10.1111/j.0956-7976.2004.00711.x> [consulté le 10 septembre 2025].

Slevc, L. R., Thompson, R. M., Salig, L. K. 2025. "Is musical ability related to second-language acquisition? A meta-analysis". *Royal Society Open Science*, 12 (1), Article 241193. [En ligne]: <https://doi.org/10.1098/rsos.241193> [consulté le 10 septembre 2025].

Xiang, C., Li, X. 2025. "Effects of mindfulness, music therapy, and social support on adolescents' test anxiety". *BMC Psychology*, 13, Article 505. [En ligne]: <https://doi.org/10.1186/s40359-025-02842-6> [consulté le 10 septembre 2025].

Zhang, Y., Ding, R., Frassinelli, D., Tuomainen, J., Klavinskis-Whiting, S., Vigliocco, G. 2023. "The role of multimodal cues in second language comprehension. *Scientific Reports*, 13, Article 20824. [En ligne]: <https://doi.org/10.1038/s41598-023-47643-2> [consulté le 10 septembre 2025].

Zheng, Y., Zhang, X., Li, J. 2025. « Passively listening to Mozart's sonata K.448 enhances verbal working memory performance : A quantitative EEG study". *Current Psychology*, 44(10), 9342-9357. [En ligne] : <https://doi.org/10.1007/s12144-025-07755-6> [consulté le 10 septembre 2025].

Zou, B., Liviero, S., Hao, M., Wei, P. 2023. Artificial intelligence and multimodal approaches in language learning: Opportunities and challenges. In: I. Levy & E. Efstathiou (Eds.), *Handbook of CALL teacher education and professional development*, Springer. p. 293-312.



© *Synergies Espagne*, n° 18, Année 2025.

Revue du GERFLINT.

ARK : <https://catalogue.bnf.fr/ark:/12148/cb426965443>

Bibliothèque nationale de France - Décembre 2025 -

Éléments sous droits d'auteur. Politique éditoriale consultable sur le site :

<https://gerflint.fr/>

<https://gerflint.fr/synergies-espagne>

