



UFR de Langues, Littératures et Civilisations Étrangères (LLCE)
Département Sciences du Langage

Master de Sciences du Langage
Parcours « Linguistique & Cognition » (LiCo)

Mémoire de Master 2

L'expérience visuelle dans le traitement du langage :
étude des corrélations phonosymboliques de forme
en condition de cécité

Lola CORCORAL

Sous la direction de
Madame Anna MARCZYK-BUKLAHA

Membres du jury :
Madame Florence BARA
Madame Florence RÉMY-EL BOUSTANI

Juillet 2025

Remerciements

J'aimerais remercier les personnes qui ont contribué à ce mémoire, sans qui ce travail n'aurait pas vu le jour. Votre aide précieuse et vos encouragements auront porté ce projet pendant deux ans.

D'abord, je remercie sincèrement mon encadrante, Anna Marczyk. Merci pour vos conseils, votre curiosité, votre énergie et votre ambition qui donnent toujours envie d'aller un peu plus loin.

Je remercie de tout cœur Mesdames Florence Bara et Florence Rémy-El Boustani d'avoir accepté de participer au jury de soutenance de ce mémoire; merci pour votre temps.

Je remercie également Monsieur Alessandro Zinna pour ses conseils lors de ma première année de master et sa gentillesse, ainsi que Raphaël Fargier, jury de première année, pour ses conseils. Merci à l'équipe pédagogique du master Lico.

Je remercie sincèrement Christophe Jouffrais et le laboratoire Cherchons Pour Voir pour leur aide dans le recrutement.

Merci évidemment à toutes mes participant-es, pour votre gentillesse, votre curiosité et votre temps. Vous n'êtes pas que des chiffres pour moi, et j'espère que vous vous amuserez à lire le fruit de votre dur labeur !

Je ne peux que me sentir chanceuse d'avoir fait ces deux années dans le master Lico avec toutes mes ami-es de promotion, dont j'espère bientôt voir le nom en haut de beaux articles. Merci à Laure, Johann et Yicheng pour m'avoir soutenue et aidé à trouver des participant-es, et Lucie, Malvina, Liem et Emma pour vos encouragements.

Evidemment, je dois remercier Amélie pour nos jolies années ensemble. Sophie et Agnès pour avoir été mes compagnes de Sciences du Langage. Merci Melizia de m'avoir toujours aidé. Certaines auront entendu parler de *bouba-kiki* en long en large et en travers...

Merci à ma famille pour toujours avoir été là pour moi.

Merci à celles et ceux que j'oublie, qui m'ont écouté et encouragé dans ce projet.

Merci encore !

Table des matières

Introduction.....	5
Chapitre 1 : Cadre théorique.....	6
1. Le phonosymbolisme : grands principes.....	6
1.1. Définition du champ d'étude.....	6
1.2. Dimensions intervenant dans l'analogie phonosymbolique.....	8
1.2.1. Dimension articulatoire.....	8
1.2.2. Dimension acoustique.....	10
1.2.3. Dimension orthographique.....	10
1.2.3. Dimension de l'expérience individuelle.....	11
1.2.4. Récapitulatif.....	12
1.2. Évaluer les correspondances phonosymboliques.....	13
1.2.1. L'étude de corpus pour évaluer le phonosymbolisme.....	13
1.2.2. L'évaluation expérimentale explicite du phonosymbolisme.....	14
1.2.3. L'évaluation expérimentale implicite du phonosymbolisme.....	15
1.3. Mécanismes d'émergence du phonosymbolisme.....	16
1.3.1. Le problème de l'émergence du phonosymbolisme : entre inné et acquis.....	16
1.3.2. Le mécanisme de co-occurrences statistiques.....	17
1.3.3. Le phonosymbolisme comme produit dérivé de l'intégration multimodale.....	18
1.3.3.1. Corrélation déficit d'intégration multimodale et déficit du phonosymbolisme.....	19
1.3.3.2. Modélisation de l'intégration multimodale et des correspondances transmodales.....	19
1.3.4. Récapitulatif : influences à la fois innées et acquises.....	21
2. Le phonosymbolisme chez les personnes déficientes visuelles.....	23
2.1. Définitions et implications de la déficience visuelle.....	23
2.1.1. Critères définitoires et types de déficiences visuelles.....	23
2.1.2. Réorganisation sensorielle et intégration multimodale chez les personnes non-voyantes.....	24
2.1.2.1. Paradigmes sensoriels dominants : implication dans l'analyse des formes.....	24
2.1.2.2. Réorganisation sensorielle : l'audio-haptique en l'absence de vision.....	25
2.1.2.3. Intégration multimodale en l'absence de vision.....	27
2.2. Implication de la vue dans le phonosymbolisme.....	28
2.3. Réponses comportementales de la population non-voyante aux correspondances transmodales.....	29
2.3.1. Correspondances transmodales simples chez les personnes déficientes visuelles.....	30
2.3.2. Effet bouba-kiki absent chez les personnes déficientes visuelles.....	31
2.1.2.3. Effet absent.....	31
2.1.2.3. Effet latent.....	32
2.3.3. Effet maluma-takete présent mais réduit chez les personnes déficientes visuelles.....	33
3. Perspectives, question de recherche.....	35
3.1. Hypothèse de la réorganisation des mécanismes phonosymboliques.....	35
3.2. Hypothèse du déficit d'intégration multimodale.....	37
3.3. Visualisation des hypothèses.....	38

Chapitre 2 : Protocole expérimental.....	39
1. Choix de la tâche : reconnaissance de paires.....	39
1.1. Généralités sur la tâche de reconnaissance de paires.....	39
1.2. Tâche de reconnaissance de paires appliquée au phonosymbolisme.....	40
1.2.1. Tâches existantes.....	40
1.2.2. Hypothèses et prédictions.....	41
2. Participants.....	42
2.1. Critères d'inclusion et d'exclusion.....	42
2.2. Recrutement.....	43
2.3. Présentation des groupes.....	43
2.3.1. Âge, genre, catégorie socio-démographique et éducation.....	43
2.3.2. Âge de perte de vue.....	44
3. Matériel.....	44
3.1. Items haptiques.....	45
3.1.1. Formes.....	45
3.1.1.1. Considérations d'exploration.....	45
3.1.1.2. Objets de forme.....	45
3.1.2. Textures.....	46
3.1.2.1. Considérations d'exploration.....	46
3.1.2.2. Objets de texture.....	46
3.2. Items sonores.....	47
3.2.1. Composition phonémique et structure des stimuli.....	47
3.2.2. Validation des logatomes via coefficient d'iconicité.....	48
3.3. Appairage logatomes et items haptiques.....	50
4. Procédure et protocole de passation.....	51
4.1. Cadre de passation.....	51
4.2. Procédure expérimentale.....	51
4.3. Entretien de fin de passation.....	53
5. Analyse des données.....	54
5.1. Variables de l'étude.....	54
5.2. Rappel des analyses et interactions testées.....	55
5.3. Pré-traitement des données.....	56
5.3.1. Modèle binomial pour l'analyse de la variance du score de mémorisation.....	56
5.3.2. Modèle binomial pour l'analyse de la variance du sentiment phonosymbolique.....	56
5.3.3. Modèle binomial pour l'analyse de la variance en fonction du type de logatome.....	57
5.3.4. Modèle linéaire pour l'analyse de la variance du temps de réaction.....	57
5.3.5. Corrélations entre variables continues.....	58
Chapitre 3 : Résultats.....	59
1. Résultats H1 – effet de congruence.....	59
1.1. Effet de congruence selon le groupe et la condition haptique.....	59
1.2. Temps de réaction au rappel selon le groupe et la condition haptique.....	60
2. Résultats H2 – effet d'apprentissage.....	61
3. Résultats H3 – corrélation mesures explicites et implicites.....	62

3.1. Sentiment phonosymbolique selon le groupe et la modalité haptique.....	62
3.2. Corrélation sentiment phonosymbolique et mémorisation.....	64
4. Résultats H4 et H5 – effet du profil acoustique des logatomes sur la tâche.....	64
4.1. Corrélation sentiment phonosymbolique et coefficient d’iconicité.....	65
4.2. Effet du type de logatome sur le sentiment phonosymbolique.....	66
4.3. Effet du type de logatome sur la mémorisation.....	66
5. Retours épilinguistiques sur la tâche.....	67
Chapitre 4 : Discussion et conclusions.....	69
1. Interprétation et discussion des résultats H1.....	69
2. Interprétation et discussion des résultats H2.....	73
3. Interprétation et discussion des résultats H3.....	74
3. Interprétation et discussion des résultats H4 & 5.....	76
4. Récapitulatif.....	77
5. Limites de l’étude.....	79
Conclusion.....	80
Bibliographie.....	83
Annexes.....	89
Annexe 1 : Profil des participants non-voyants.....	89
Annexe 2 : Profil des participants voyants.....	89
Annexe 3 : Paires de forme.....	90
Annexe 4 : Paires de texture.....	91
Annexe 5 : Questionnaire participant.....	92
<i>Déclaration sur l’honneur de non-plagiat.....</i>	<i>94</i>

Introduction

Le phonosymbolisme renvoie au champ d'étude s'intéressant à l'habileté humaine à associer certains phonèmes à des impressions synesthésiques, telles que des qualités de forme, de taille ou encore de texture. L'exemple le plus connu de corrélation phonosymbolique est l'association des consonnes sonantes (par exemple, /m/) avec le trait de {rondeur}, et des consonnes occlusives sourdes avec le trait {angulaire}, communément connu sous le nom d'*effet maluma-takete* (Köhler, 1929). L'origine de ces associations demeure controversée (p ex. Sidhu & Pexman, 2018). Elles font néanmoins l'objet d'un intérêt particulier en sciences cognitives, car ces associations étonnantes semblent innées (p ex. Ozturk *et al.*, 2013) et perçues à l'échelle de l'humanité (Ćwiek *et al.*, 2021).

Comme pour tout phénomène cognitif, le mode de traitement typique n'est jamais mieux compris qu'en observant un mode de traitement dit atypique. Etant donné que le phonosymbolisme est intimement lié à l'expérience sensorielle, nous avons interrogé ses origines à travers l'étude comportementale des personnes en contexte de privation sensorielle, ici les personnes non-voyantes. Une littérature récente existe sur ce sujet et rapporte des résultats hétérogènes ainsi que des comportements particulièrement atypiques chez cette population.

Dans cette étude, nous souhaitons ainsi faire une analyse comparative des tendances phonosymboliques des personnes voyantes et non-voyantes, pour aborder les questions suivantes : pourquoi les personnes déficientes visuelles expriment-elles des tendances phonosymboliques atypiques ? La vue est-elle nécessaire à l'émergence de corrélations phonosymboliques typiques ?

Dans un premier chapitre, nous définirons notre champ d'étude. Nous mettrons d'abord en lumière les caractéristiques du profil phonosymbolique typique ainsi que les méthodes utilisées pour le mesurer. Nous aborderons ensuite la question de la déficience visuelle et de ses implications sensorielles, avant de décrire le profil phonosymbolique des personnes non-voyantes. Cette discussion nous mènera à cibler notre question de recherche et nos principales hypothèses. Dans notre second chapitre, nous décrirons le protocole expérimental qui nous a permis de les interroger. Ce projet, mis en place dès avril 2025, a pris la forme d'une tâche de mémorisation haptique de paires congruentes et incongruentes au niveau phonosymbolique. Nous finirons cette partie en présentant le prétraitement des données de l'étude, en vue de la présentation des résultats, au troisième chapitre. Nous finirons notre développement, dans le quatrième chapitre, en discutant ces résultats.

Chapitre 1 : Cadre théorique

1. Le phonosymbolisme : grands principes

1.1. Définition du champ d'étude

Le symbolisme phonétique, ou phonosymbolisme, est le champ d'étude qui s'intéresse aux correspondances transmodales systématiques entre des sons et des représentations non-acoustiques. Autrement dit, ce domaine explore les cas où des sons constitutifs du signifiant viennent à revêtir des propriétés sémantiques : on interroge donc l'expressivité des phonèmes. L'objet le mieux documenté du phonosymbolisme sont les corrélations de formes, soit l'association systématique entre un phonème et une qualité de forme, généralement l'opposition de qualité {rond/angulaire}, communément connu sous le nom d'*effet maluma-takete* (Köhler, 1929) ou *effet bouba-kiki* (Ramachandran & Hubbard, 2001). Ainsi, nous décrirons surtout celles-ci.

Il est généralement admis, depuis la publication fondamentale du *Cours de Linguistique Générale* (de Saussure, 1916), que le signe linguistique est arbitraire. Autrement dit, pour la majorité des signes composant le code des langues, les *signifiants* (les formes phoniques des mots) n'entretiennent pas de liens naturels avec leurs *signifiés* (les concepts associés). Ce principe, communément appelé *l'arbitraire du signe*, est aujourd'hui nuancé par les apports édifiants du symbolisme phonétique, qui se définit comme l'étude des phénomènes linguistiques de lien *motivé* entre un phonème et une représentation non-acoustique. On a par exemple remarqué que, si l'on demande à un grand groupe de participants d'associer les pseudo-mots /maluma/ et /takete/ à une forme géométrique {ronde} ou {angulaire}, ils associent de façon quasi-systématique (à hauteur de 95% de réussite) le pseudo-mot /maluma/ à la forme {ronde} et /takete/ à la forme {angulaire} (Köhler, 1929). A priori, ni /maluma/ ni /takete/ n'ont de signifié : c'est donc bien que ces signifiants sont déjà à eux seuls porteurs d'un sémantisme et sont donc motivés. On envisage un signe phonosymbolique comme un signe iconique, dans la mesure où la forme du référent ressemble à et contraint dans une certaine mesure la forme du signifiant. On dira que là où il y a icône, le signifiant et le référent d'un signe partagent donc des traits communs, généralement des qualités perceptuelles. Jakobson (1980 : 223) parle d'*analogie intersensorielle*. L'iconicité d'un signe peut être très forte comme elle peut être plus faible, ce qui nous amène à l'envisager comme un continuum. Un signe présente donc un *degré d'iconicité* qui peut tout aussi bien être élevé, comme c'est le cas pour les stimuli /maluma/ et /takete/ de Köhler (1929), que nul.

Les corrélations phonosymboliques sont étudiées expérimentalement à partir du 20^e siècle, mais la question de la motivation linguistique est loin d'être une nouveauté et est à minima vieille de deux-mille ans. Jakobson (1965) évoque ainsi le très ancien Cratyle de Platon, où la question du phonosymbolisme est déjà évoquée en considérant que les lettres et phones sont “*l'imitation de l'essence*” (§ 425c) de leurs signifiés et référents.

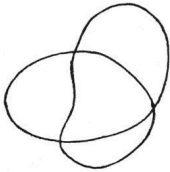
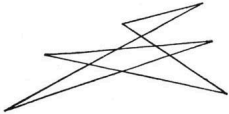
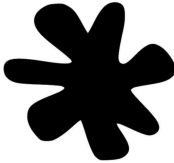

Köhler (1929)		Ramachandran & Hubbard (2001)	
/maluma/	/takete/	/buba/	/kiki/
			

Figure 1. Stimuli visuels pour /maluma/ et /takete/ de Köhler (1929) puis /buba/ et /kiki/ de Ramachandran & Hubbard (2001)

En cela, il est considéré que le phonosymbolisme est un phénomène de *correspondance transmodale*, c'est-à-dire un effet d'analogie entre deux stimuli de modalités sensorielles différentes, où la perception d'un stimulus induit une association intuitive avec un percept d'une autre modalité sensorielle. Spence (2011) définit une correspondance transmodale comme un phénomène à trois propriétés :

- (1) Les deux indices sensoriels associés doivent être, dans le monde réel, **complémentaires**. C'est-à-dire, ils doivent renseigner sur le même objet du monde. Par exemple, la perception d'une fréquence {aiguë} et d'une taille {petite} synchronisées spatio-temporellement proviennent généralement du même objet du monde. Spence parle d'effet de compatibilité entre les signaux (2011 : 973).
- (2) La correspondance doit opérer entre deux extrêmes des deux stimuli, soit, elle doit être polarisée. La correspondance repose donc sur un effet d'opposition entre deux aspects d'un même continuum sensoriel. Dans le cas des correspondances de forme, on polarise ainsi l'expression d'une qualité de forme et qualité acoustique; un pôle de ce continuum est compatible avec l'un et pas avec l'autre, tel que {rond} est compatible avec le critère acoustique [+continu] et {angulaire} avec [-continu] (p. ex Monaghan *et al.*, 2012).
- (3) Enfin, la correspondance doit s'exprimer de façon analogue chez une proportion significative de la population, c'est-à-dire être quasi-**universelle**. Dans le cadre du phonosymbolisme de forme, l'universalité des correspondances du type *bouba-kiki* a été observée à maintes reprises (par ex. Davis, 1961; Ćwiek *et al.*, 2021), pointant effectivement vers une correspondance partagée au niveau de l'humanité.

Ces associations systématiques entre sons et sens ont été décrites dans différentes modalités perceptuelles. A ce titre, Boudier (2018) recense sept grandes catégories de corrélations phonosymboliques, dont le phonosymbolisme de forme (p. ex. Köhler, 1929). Pour en citer d'autres, le phonosymbolisme des sensations tactiles (p. ex. Etzi *et al.*, 2015) étudie les associations phonémiques avec les sensations {lisses~rugueuses} ou {dur~doux}. Le phonosymbolisme de taille (p. ex. Sapir, 1929) étudie, lui, l'opposition {petit~grand}.

Bien que ces corrélations aient leurs propres caractéristiques, il arrive régulièrement qu'un phonème évoque plusieurs impressions. C'est par exemple le cas du /i/, qui peut revêtir une grande diversité d'impressions selon le contexte. Il est par exemple systématiquement associé aux traits {petit}, {angulaire}, {dur}, {lumineux}, etc (French, 1977). Cet effet est appelé la pluripotentialité : un même phonème est potentiellement associé à un grand volet de qualités sémantiques latentes (d'associations intersensorielles) qu'on peut faire émerger selon le contexte (Chastaing, 1964; Winter *et al.*, 2019). Sur la base de leurs propres études, Westbury *et al.* (2018 : 147) suggèrent "*qu'il est possible [...] qu'un petit nombre d'associations [...] sous-tendent la majorité des associations phonosymboliques*"¹. Il est donc intéressant, dans l'étude même du phonosymbolisme entre un son et une qualité géométrique, de ne pas se restreindre à étudier le phonosymbolisme de forme.

1.2. Dimensions intervenant dans l'analogie phonosymbolique

Nous venons de voir que le phonosymbolisme se base sur une relation d'analogie entre deux signaux. Il convient donc de se demander quels sont les vecteurs de cette analogie, quelles sont les dimensions intervenant dans le degré d'iconicité d'un mot ?

Dans cette partie, nous aborderons quatre grandes dimensions qui semblent intervenir dans l'orientation des corrélations phonosymboliques : la dimension articulatoire, acoustique, la graphie et finalement, l'expérience culturelle.

1.2.1. Dimension articulatoire

La dimension la plus lourde dans le degré d'iconicité semble être, de façon assez consensuelle dans la littérature, la dimension articulatoire des phonèmes. L'association entre un phonème et un trait sémantique serait médiée par la proprioception du mouvement articulatoire, dans le sens où l'on observerait un effet d'analogie entre un trait articulatoire et un trait sémantique, ce que Newman (1933) appelle le *facteur kinésique*. L'hypothèse de

¹ Dans l'original : "it is possible [...] that a small number of basic associations [...] underlie the majority of sound symbolic associations".

l'analogie proprioceptive est intéressante dans la mesure où elle s'inscrit dans la lignée de la théorie motrice de la perception de la parole (Lieberman *et al.*, 1985) qui défend que l'unité fondamentale de la perception de la parole est le geste moteur articulatoire, non pas les composantes acoustiques des sons.

Selon le postulat de l'analogie kinesthésique, le sentiment de percevoir une voyelle comme étant {petite} ou {large} serait une mise en corrélation directe avec son degré d'aperture. En ce sens, la voyelle fermée /i/ est perçue comme {petite} car l'espace entre la langue et le palais est refermé lors de son articulation, alors qu'à l'opposé la voyelle ouverte /a/ est perçue comme {large} car son aperture articulatoire est large (Newman, 1933). Il est fréquemment proposé que l'arrondissement des lèvres induit également la perception de certains phonèmes comme {ronds} (p ex. D'Onofrio, 2014).

Quant aux consonnes, Nobile (2015) montre que le [voisement] prédit dans 91% des cas le sentiment phonosymbolique de {rondeur}, l'inverse étant vrai pour une consonne sourde. Westbury (2005) observe, quant à lui, une corrélation entre les consonnes continues (sonantes, glides et fricatives) et le trait sémantique {rond}, ainsi qu'une corrélation entre les consonnes occlusives et le pôle {angulaire}. Ici, la proprioception d'un mode articulatoire souple et continu semble permettre l'analogie vers une forme {ronde}. A l'inverse, la proprioception d'un mode articulatoire bref et saccadé semble nourrir les sentiments phonosymboliques {angulaires}. McCormick *et al.* (2015) trouvent un résultat similaire.

L'hypothèse de l'analogie articulatoire a tout de même ses limites. Bottini *et al.* (2018) suggèrent que, si le phonosymbolisme de la forme était exclusivement modulé par des analogies motrices, les différents phonèmes ci-dessous devraient être classés ainsi sur le continuum d'iconicité {rond~angulaire}, en fonction du voisement et mode articulatoire :



	SONNANTE	FRICATIVE VOISÉE	FRICATIVE SOURDE	OCCLUSIVE VOISÉE	OCCLUSIVE SOURDE	
{rond}	/n/ /m/ /l/	/v/ /z/ /ʒ/	/f/ /s/ /ʃ/	/b/ /d/ /g/	/t/ /k/ /p/	{angulaire}

Figure 2. Échelle hypothétique des prédicteurs pour {rond~angulaire} respectant la dimension articulatoire, selon Bottini *et al.* (2018)

Or, cette prédiction n'est pas analogue à des résultats effectifs. Le trio /v/, /z/ et /ʒ/ est, par exemple, généralement perçu comme étant plus {rond} que /f/ /s/ /ʃ/ (p. ex. McCormick *et al.*, 2015). La dimension articulatoire, bien que pertinente, n'est donc pas suffisante pour expliquer les configurations finales des associations phonosymboliques.

1.2.2. Dimension acoustique

Un autre facteur d'analogie semble être la dimension acoustique des phonèmes. Fort & Schwartz (2022) suggèrent par exemple que la balance spectrale d'un phonème et sa continuité sont les propriétés les plus adeptes de prédire les associations phonosymboliques, où une balance spectrale concentrée dans les fréquences graves et une répartition continue du signal est généralement associée à la qualité de forme {rond}. La répartition formantique semble également être une dimension idoine pour prédire et classer le phonosymbolisme des voyelles (Knoeferle *et al.*, 2017). Sur le plan fréquentiel, de nombreuses études montrent en effet que les corrélations transmodales sont perçues même en l'absence de matériaux linguistiques, notamment à l'écoute de sons purs opposés en fréquence moyenne, aiguë ou grave (p ex. O'Boyle & Tarte, 1980; Hamilton-Fletcher *et al.*, 2018;). Au même titre, Parise & Spence (2012) ont mis en évidence que les ondes apériodiques (caractérisées par du bruit) sont significativement corrélées au trait {angulaire}, alors que des ondes périodiques (plus agréables à l'écoute) sont corrélées au trait {rond}.

Ici, on pourrait avancer que les propriétés acoustiques émergent de configurations articulatoires et que, par conséquent, les deux dimensions sont liées. Cela étant dit, l'avantage d'étudier la dimension acoustique est qu'elle est plus facilement quantifiable et permet ainsi des analyses corrélationnelles précises. Contrairement à la phonation, la dimension acoustique ne se limite pas à une modélisation en matrices binaires et permet, entre-autre, de prendre facilement en compte les nuances dans les ressentis phonosymboliques de la population, notamment en quantifiant un degré d'iconicité comme un continuum. Cela pourrait notamment permettre de prendre en compte les cas d'un sentiment phonosymbolique qui n'est pas tranché. En partie pour ces raisons, c'est la dimension acoustique qui nous intéressera le plus dans notre étude.

1.2.3. Dimension orthographique

En troisième lieu, il semble que la graphie influence le degré d'iconicité de certaines unités. Une étude récente de Bottini *et al.* (2018) s'est intéressée à l'influence de l'orthographe dans la mise en place du paradigme *maluma-takete*. Cette étude pose l'observation initiale comme quoi les lettres de l'alphabet romain sont elles-même des formes plus ou moins rondes ou angulaires. Comme évoqué plus tôt, étant donné que des analogies d'ordre strictement articulatoires entre les sons et les traits {rond~angulaire} ne permettent pas de rendre compte de la diversité des corrélations phonosymboliques, Bottini *et al.* (2018) proposent l'hypothèse BOAH (*Blending Orthography and Articulatory Hypothesis*), selon laquelle la forme des lettres s'implémente par dessus l'iconicité kinésique et modifie ainsi

l'orientation des corrélations phonosymboliques. Leurs résultats soutiennent leur hypothèse, particulièrement pour les occlusives voisées /b, d, g/ qui, bien qu'elles soient [-continues] et donc potentiellement {angulaires}, sont très fortement corrélées au trait {rond}. Ce fait s'explique parce qu'elles sont [voisées], mais également car leurs graphèmes <b, d, g> sont graphiquement ronds. La forme des lettres influence donc le phonosymbolisme des mots. Particulièrement, les locuteurs d'une langue réinvestissant l'alphabet latin semblent présenter des effets phonosymboliques légèrement plus forts que les autres (Ćwiek *et al.*, 2021).



Figure 3. Illustration des propriétés géométriques partagées entre des lettres et des formes, issu de Cuskley *et al.* (2017 : 121)

1.2.3. Dimension de l'expérience individuelle

Enfin, des expériences sensorielles d'ordre individuel et culturel semblent également intervenir dans l'iconicité d'un mot. L'étude de Bremner *et al.* (2013) décrit par exemple une corrélation phonosymbolique de goût complètement contraire aux prédictions que l'on aurait pu faire sur la base des dimensions articulatoires, acoustiques ou orthographiques seules. Dans leur étude, il a été demandé à des participants de Namibie d'associer un goût d'eau {gazeuse} ou {plate} aux formes angulaires et rondes. Étonnamment, les participants associent {gazeuse} à /buba/ et {plate} à /kiki/ là où la plupart des études antérieures retrouvent l'effet inverse. Selon les chercheurs, ces résultats seraient dûs au fait que les marques locales d'eau gazeuse s'appuient sur une iconographie médiatique très visuellement {ronde}. En cela, l'influence de la modalité visuelle dans l'effet bouba-kiki ne serait donc pas limitée aux facteurs strictement linguistique (l'orthographe et l'articulation), mais s'étend potentiellement à toutes les récurrences systématiques entre une forme visuelle et une autre modalité sensorielle rencontré par un individu au cours de sa vie. La culture et l'expérience individuelle ont donc un impact dans le degré d'iconicité des mots.

1.2.4. Récapitulatif

A l'issue de cette section, nous avons pu voir que le facteur d'analogie phonosymbolique n'est pas facilement identifiable et n'est probablement pas unique mais à la croisée des multiples dimensions explorées plus tôt. De grandes tendances phonosymboliques émergent malgré tout à travers la population. Afin d'en rendre compte, nous avons recueilli les résultats de douze études évaluant la polarisation phonosymbolique de différents phonèmes. Ce tableau permet d'appréhender, concrètement, quelles unités phonologiques semblent les plus {rondes} ou les plus {angulaires}. Les résultats ne sont pas toujours homogènes, néanmoins, à la croisée de toutes les études, les phonèmes /b/ /l/ /m/ /n/ /w/ /u/ /o/ et /ɔ/ sont toujours représentatifs du pôle {rond} et les phonèmes /k/ /t/ /e/ /a/ /i/ /ɛ/ et /i/ toujours représentatifs du pôle {angulaire}. Ci-dessous le tableau :



Études	 {rond}	 {angulaire}
Cuskley <i>et al.</i> (2017)	[d] [g] [s] [f]	[z] [v] [t] [k]
D'Onofrio (2014)	[+cons] [+labial] [+sonore] [+vocalique] [+postérieur]	[+cons] [+alvéolaire] [+vélaire] [-sonore] [+vocalique] [+antérieur]
Fort <i>et al.</i> (2015)	[b] [ʃ] [d] [l] [m] [n] [s] [p] [ʒ] [g] [nasal] [o] [u]	[f] [k] [t] [v] [z] [i] [e]
Kovic <i>et al.</i> (2010)	[+labial]	[-labial]
Maurer <i>et al.</i> (2006)	[a] [u]	[i] [e] [ʌ]
McCormick <i>et al.</i> (2015)	[vocalique] [+labial] [m] [l]	[vocalique] [-labial] [p] [t] [k] [+affriqué]
Monaghan <i>et al.</i> (2012)	[+cons] [+continu] [+vocalique] [+ouvert] [+postérieur]	[+cons] [-continu] [+vocalique] [+fermé] [+antérieur]
Nielsen & Rendall (2011)	[b] [m] [l] [g]	[p] [t] [k]
Nobile (2015)	[+cons] [+sonore] [-continu] [+cons] [+sonore] [+continu] [+cons] [+palato-vélaire] [+nasal] [+vocalique] [+labial] [+vélaire]	[+cons] [-sonore] [-continu] [+cons] [-sonore] [+continu] [+cons] [+aleo-dental] [+oral] [+vocalique] [-labial] [+palatal]
Sidhu <i>et al.</i> (2021)	/u/ /m/ /oo/ /b/	/ai/ /tʃ/ /k/ /ʃ/ /ɜ:/ /r/ /t/ /s/ /l/ unités lexicales longues
Westbury (2005)	[+cons] [+continu]	[+cons] [-continu]
Westbury <i>et al.</i> (2018)	<o> <u> /ow/ /uw/ [lateral approximant] /b/ /ɔ/ /u:/ /ɒ/ <m>	<x> <k> <c> <t> /k/ [labiodental] <k> <t> <am> <bl> /t/ /k/

Tableau 1. Phonèmes et traits représentatifs des pôles {rond} et {angulaire} selon douze études.

1.2. Évaluer les correspondances phonosymboliques

A l'origine, l'étude du phonosymbolisme s'est basée sur l'intuition des locuteurs intéressés. Platon proposait déjà dans *Le Cratyle*, sur la base de sa propre intuition, plusieurs corrélations aujourd'hui bien établies, par exemple une interprétation du /i/, qui est "*fin, subtil, et capable de pénétrer les autres choses*" (§ 427e). Si le sentiment de la langue a été moteur dans l'étude du symbolisme phonétique, cette méthode n'est plus majoritaire.

Dans cette partie, nous résumons les méthodes d'exploration des corrélations phonosymboliques moderne les plus courantes, mises en place pour aborder le phonosymbolisme de façon non-subjective, à savoir :

- l'étude de corpus ;
- les tâches expérimentales explicites ;
- et les tâches expérimentales implicites.

1.2.1. L'étude de corpus pour évaluer le phonosymbolisme

L'étude scientifique du phonosymbolisme a débuté par des études de corpus sur des langues naturelles. L'hypothèse principale de ce type d'étude est que, plus un son est représenté dans un champ lexical incluant un sème particulier, plus ce son connote ce sens. Par extension, ce son serait l'imitation de ce sème. Par exemple, Thorndike (1945) s'est intéressé au phonosymbolisme de la voyelle /i/. Il a donc listé tous les mots monosyllabiques de l'anglais contenant en noyau les voyelles [i] ou [ɪ] et ayant une fréquence lexicale supérieure ou égale à une occurrence tous les million de mots. A partir de ce corpus de 527 mots, Thorndike a compté combien sont liés à la notion de {petitesse}. Ses calculs montrent que la probabilité qu'un mot contenant les voyelles [i] ou [ɪ] connote la notion {petitesse} en anglais est douze fois plus probable que pour un mot contenant les voyelles [u] et [o] par exemple. Sur la base de ses résultats, Thorndike suggère qu'il y a corrélation phonosymbolique entre [i] et le trait {petit} en anglais.

Des études sur corpus sont toujours faites aujourd'hui. A titre d'exemple, Winter *et al.* (2022) ont démontré que les rhotiques apparaissent de façon significativement plus fréquente dans des mots désignant des sensations rugueuses que des sensations douces ou lisses. Cela a été démontré pour l'anglais, dans 38 langues Indo-Européennes et 331 langues du monde issues de 84 familles linguistiques différentes.

On remarquera ici que les études de corpus permettent d'étudier l'universalité des corrélations phonosymboliques. Il faut tout de même préciser que cette méthodologie mesure la **systematicité** de la relation entre un son et un sème, non son iconicité.

1.2.2. L'évaluation expérimentale explicite du phonosymbolisme

L'évaluation expérimentale explicite du phonosymbolisme réfère à un ensemble de tests perceptifs permettant de récolter une mesure introspective du phonosymbolisme à grande échelle et constitue aujourd'hui la majorité des expérimentations mises en place pour évaluer les corrélations phonosymboliques (Motamedi *et al.*, 2019). Ces tâches demandent aux participants d'exprimer une association consciente entre (au moins) deux items, nécessitant donc une réflexion intentionnelle quant aux items présentés. L'hypothèse récurrente derrière ce type de tâche est qu'un mot ou une unité sub-lexicale est phonosymbolique si elle guide et facilite une association de façon significative (Motamedi *et al.*, 2019). Plus une association est partagée par l'ensemble des participants, plus la corrélation phonosymbolique est forte.

Parmi les possibles, la tâche d'association en choix contraint, introduite par ailleurs par Köhler (1929) et Sapir (1929), est la plus utilisée. 42 sur les 99 expériences phonosymboliques recensées par Westbury *et al.* (2018) appliquent cette tâche. Lors de cette épreuve, on donne un ou deux logatomes au participant afin qu'il les associe à d'autres types d'items, par exemple des formes visuelles. Cette tâche permettrait de mettre en évidence une corrélation générale ou, à mesure répétée, l'expressivité propre de chaque phonème. Les tâches expérimentales explicites restent les tâches expérimentales les plus simples pour mettre en évidence des tendances phonosymboliques chez un groupe d'individus. Néanmoins, ces tâches présentent deux grandes limites :

- (1) D'abord, il est difficile d'estimer si les phonèmes qu'un participant associe aux pôles {angulaire~rond} sont des *vrais positifs* (Westbury *et al.*, 2018 : 123) ou simplement associé par opposition avec le trait non-congruent. Par exemple, comment savoir si les dimensions intervenant dans le degré d'iconicité de forme pour /buba/ sont prédicteurs pour {rond} ou {non-angulaire} ? Dans des contextes où on présente simultanément deux logatomes à associer à deux formes, on ne sait pas vraiment si on évalue la valeur phonosymbolique réelle de chaque logatome ou les relations qui émergent d'effets d'opposition entre les quatre items.
- (2) Ensuite, Nielsen & Rendall (2012 : 117) suggèrent que les tâches d'association en choix contraint variantes offrent des résultats sur-représentatifs de la puissance des corrélations phonosymboliques réelles. Ceux-ci suggèrent effectivement que ces tâches sont "*transparentes*" et mettent en exergue des oppositions sur lesquelles le participant peut facilement s'appuyer. De plus, l'opposition phonologique est généralement maximisée pour que le contraste permette l'observation de l'effet (Westbury *et al.*, 2018). Les tâches d'association induisent donc en partie des tendances phonosymboliques.

1.2.3. L'évaluation expérimentale implicite du phonosymbolisme

Nous venons de voir que la méthode de description expérimentale explicite est la plus populaire, mais présente des limites importantes. Il convient donc de se demander si les corrélations phonosymboliques, dans notre cas des associations forme-son, émergent également hors du contexte expérimental explicite. Pour répondre à ces questions, nous aborderons les méthodes d'évaluation implicite du phonosymbolisme.

L'évaluation expérimentale implicite du phonosymbolisme permet de récolter une mesure non-introspective du phonosymbolisme. On mesure ici des données dérivées de l'association phonosymbolique qui devrait, selon les hypothèses associées à des tâches implicites, offrir un type d'avantage cognitif. Ces tâches se basent sur des effets de congruence, et les mesures dérivées peuvent être, par exemple, des temps de réaction ou des scores de mémorisation. Les tâches axées sur l'appréhension implicite de phonosymbolisme sont basées sur la notion de congruence versus incongruence entre deux stimuli (ici, entre la forme phonologique et la forme visuelle présentée aux participants). On considère que deux items sont congruents quand ils se potentialisent dans des conditions de traitement cognitif. L'hypothèse ici suppose que, si les associations phonosymboliques sont conceptualisées à un haut niveau de traitement, un participant manipulant des relations congruentes au niveau phonosymbolique dans un tâche de mémorisation ou autre devrait bénéficier d'un avantage par rapport à des relations incongruentes. Cette hypothèse a été validée par plusieurs études, mais il est bon de remarquer que les effets du phonosymbolisme sont beaucoup plus faibles lorsqu'on les étudie de façon implicite.

L'avantage phonosymbolique a été démontré dans une tâche de décision lexicale (Westbury, 2005), un test d'association implicite (Parise & Spence, 2012), en condition d'amorçage perceptif (Sidhu & Pexman, 2017) et même à travers des comparaisons de N400 en EEG (Kovic *et al.*, 2010). Notre étude s'intéresse plus particulièrement à l'avantage phonosymbolique en tâche de mémorisation (par exemple Aveyard, 2012; Nielsen & Rendall, 2012). Une tâche de mémorisation sur critère phonosymbolique pose l'hypothèse suivante : si la congruence facilite la mémorisation d'une paire forme/logatome, alors le phonosymbolisme est une habileté naturelle de haut niveau cognitif. Dans cette idée, Aveyard (2012) a mis en place une tâche de mémorisation avec des paires congruentes et incongruentes au niveau phonosymbolique. Dès les premier tour, l'étude montre que les paires congruentes sont significativement mieux mémorisées que celles incongruentes. Cela suggère que le phonosymbolisme facilite la mémorisation, bien que l'effet soit faible.

L'avantage de cette méthode provient du fait que les données comportementales recueillies sont le reflet de la robustesse des réseaux associatifs encodant les corrélations phonosymboliques, peu importe le degré de conscientisation de ces associations. Ces données sont ainsi moins sensibles aux introspections (parfois trompeuses) des participants.

1.3. Mécanismes d'émergence du phonosymbolisme

Nous avons vu que les associations phonosymboliques reposent sur des mécanismes analogiques dont le vecteur pourrait provenir de différents domaines (par ex. articulatoire ou acoustique). Ces corrélations ont été testées à maintes reprises de façon explicite et implicite, démontrant qu'elles sont systématiques, robustes et partagées par un grand nombre de la population. Désormais, reste la question : d'où émergent ces associations ? Sont-elles innées ou acquises ? Dans cette partie, nous aborderons le problème de l'origine cognitive du phonosymbolisme en présentant les hypothèses phares de la littérature.

1.3.1. Le problème de l'émergence du phonosymbolisme : entre inné et acquis

La question de l'origine des tendances phonosymboliques est une problématique controversée et il n'y a, à ce jour, pas de consensus global quant aux mécanismes qui sous-tendraient le phonosymbolisme. Le débat concerne notamment l'aspect inné ou acquis des corrélations : d'un côté, ces corrélations semblent innées, car elles sont présentes chez des enfants d'un très jeune âge (par ex. Ozturk *et al.*, 2013, Maurer *et al.*, 2006) et semblent être partagées par une très grande proportion de l'humanité (Ćwiek *et al.*, 2021), ce qui mène même certain-es chercheur-ses à postuler que les phénomènes comme l'*effet bouba-kiki* ou *maluma-takete* sont universels et reflètent une propriété fondamentale des langues, aux origines du langage (par ex. Ramachandran & Hubbard, 2001).

D'autres études indiquent plutôt que la perception de ces corrélations est acquise : il a par exemple été montré que, même si les enfants expriment les corrélations dès un très jeune âge, elles sont assez faibles à l'origine et deviennent plus robustes avec l'âge (Fort *et al.*, 2018). Comme évoqué précédemment, il a également été montré que certaines corrélations phonosymboliques sont affectées par des phénomènes culturels comme l'iconographie des médias locaux (Bremner *et al.*, 2013) ou de l'écriture (Bottini *et al.*, 2018; Ćwiek *et al.*, 2021), ce qui pointe bien vers un aspect acquis de ces corrélations.

La littérature déborde d'hypothèses quant à quoi serait dû le phonosymbolisme. Sidhu et Pexman (2018) en recensent cinq grandes catégories, dont celle qui nous intéressera le plus ici : le mécanisme d'intégration des cooccurrences statistiques de stimuli sensoriels dans l'environnement. En effet, et bien que plusieurs mécanismes semblent interagir et affecter la configuration finale des correspondances phonosymboliques, la littérature autour de ce champ d'étude considère que le mécanisme de cooccurrence statistique sous-tend la majorité des tendances observables (notamment l'*effet maluma-takete*). Dans les deux prochaines parties, nous essayerons de montrer en quoi ce mécanisme est particulièrement intéressant, dans la mesure où il permet à la fois de justifier l'impression d'universalité des effets et les variations inter-individuelles observables dans ce champ d'étude.

1.3.2. Le mécanisme de co-occurrences statistiques

Nous venons de voir que les corrélations phonosymboliques semblent universelles, dans le sens où elles émergent à priori de la même façon peu importe la population questionnée (Ćwiek *et al.*, 2021; Ćwiek *et al.*, 2024). Une partie de la littérature s'intéressant au phonosymbolisme suggère que cette universalité serait dû au fait que le mécanisme d'émergence des effets du type *maluma-takete* s'appuie sur les propriétés physiques physiques des objets du monde. Bien que les mécanismes d'émergence des correspondances transmodales soient encore controversés, il est généralement accepté qu'elles sont le résultat d'une exposition systématique à des associations d'indices perceptifs disponibles dans l'environnement. L'émergence d'une relation analogique entre un son et une qualité de forme serait basée sur l'intégration des *cooccurrences statistiques* de stimuli dans l'environnement.

Autrement dit, le phonosymbolisme serait une extension, ou un produit dérivé des correspondances transmodales simples acquises avec le temps. Par expérience, on apprend par exemple qu'un petit objet émet des sons aigus; en retour un son aigu va donc évoquer un petit objet via un mécanisme associatif voire prédictif. Ce type de configuration fréquentielle se retrouve par exemple dans le phonème /i/, typiquement associé avec le trait {petit} (par ex. Fort *et al.*, 2015) ou avec les fricatives comme /s/, dont l'énergie est concentrée dans les hautes fréquences. Ces associations seraient dues à la généralisation de connaissances issues de l'environnement : les petites choses ont des fréquences de résonance élevées, l'inverse étant vrai pour des objets larges (Gallace & Spence, 2006). Spence (2011) nomme ce type d'association des correspondances transmodales structurales.

Fort & Schwartz (2022) suggèrent ainsi que le phonosymbolisme est un phénomène intrinsèquement lié aux correspondances transmodales non-linguistiques. Selon ces deux chercheurs, des critères acoustiques seraient à la source des phénomènes phonosymboliques et ceux-ci peuvent être expliqués grâce à la configuration spectrale d'un phonème, qui évoquent, par analogie, les configurations spectrales d'autres signaux de l'environnement.

Ces chercheurs ne sont pas les seuls à avoir défendu cette hypothèse (voir par ex. le *frequency code* d'Ohala, 1994; French, 1977; ou Parise & Spence, 2012). De façon intéressante pour nous, Fort & Schwartz (2022) défendent par ailleurs que la corrélation entre une forme {ronde} et les basses fréquences est naturellement présente dans l'environnement, notamment car la fréquence de résonance moyenne d'un objet est prévisible par sa forme. Selon leurs investigations personnelles et antérieures, il a déjà été prouvé mathématiquement que les objets aux formes rondes ont des fréquences de résonance moyennes plus basses que les objets aux formes angulaires (Kac, 1966). La population typique a par ailleurs intégré

cette différence et peut, par exemple, correctement inférer la forme {ronde} ou {angulaire} d'un objet à partir des signaux naturels qu'il émet (Fort & Schwartz, 2022 : 5).

Enfin, si l'expressivité des phonèmes émerge d'équivalences intersensorielles médiées par des propriétés acoustiques, en partie des jeux de fréquence (/i/ exploite des éléments acoustiques analogues aux bruits émis par une forme {angulaire}), cela revient à dire qu'il y a une projection des correspondances transmodales non-linguistiques sur les corrélations du type *maluma-takete*. Cette hypothèse est intéressante dans la mesure où elle permet d'expliquer l'universalité apparente du phonosymbolisme, qui se baserait donc sur une généralisation des propriétés physiques de l'environnement.

Dans notre prochaine partie, nous verrons en quoi cette hypothèse permet également de rendre compte des variations interindividuelles dans les tendances phonosymboliques, en suggérant notamment que les correspondances transmodales sont issues du mécanisme d'intégration multimodale.

1.3.3. Le phonosymbolisme comme produit dérivé de l'intégration multimodale

Nous venons de voir que le phonosymbolisme semble être une projection de correspondances transmodales non-linguistiques (entre des signaux sonores et des percepts non-sonores) sur les phonèmes constitutifs des langues. Ces correspondances transmodales sont à priori universelles, car elles s'appuient sur les propriétés physiques de l'environnement. Nous pourrions désormais nous demander : d'où proviennent ces correspondances transmodales ?

Spence (2011), expert des correspondances transmodales gustatives, suggère que les correspondances transmodales sont un produit dérivé du processus perceptif d'*intégration multimodale*. L'intégration multimodale renvoie aux situations où deux signaux sensoriels redondants (qui proviennent potentiellement de la même source au même moment) sont unis dans la perception afin de percevoir un tout cohérent. Lorsque deux signaux sont intégrés, leur traitement est facilité et optimisé. Cet effet se traduit généralement en un traitement plus rapide et précis de l'information. En contre-point, l'intégration induit une analyse distordue du détail des signaux individuels (l'interprétation unimodale), qui sera moins accessible aux individus percevants (Ernst, 2005). L'intégration audio-visuelle est un des sujets phares de ce domaine. Par exemple, il est aujourd'hui admis que la vue affecte la perception des phonèmes, soit l'*effet McGurk* (McGurk & MacDonald, 1976), ou la localisation du son dans l'espace, dit l'*effet ventriloque* (Howard & Templeton, 1966).

1.3.3.1. Corrélation déficit d'intégration multimodale et déficit du phonosymbolisme

Soutenant l'hypothèse comme quoi le phonosymbolisme émerge du mécanisme d'intégration multimodale, il a été remarqué que pour un même individu, un déficit d'intégration multimodale semble être corrélé à une diminution des tendances phonosymboliques. Par exemple, il a été montré à plusieurs reprises que les adultes sur le spectre autistique présentent des corrélations phonosymboliques de forme moins robustes que des participants typiques appariés sur l'âge chronologique (Oberman & Ramachandran, 2008; Occelli *et al.*, 2013). Dans une tâche d'association en choix contraint classique du paradigme *bouba-kiki*, le groupe TSA présente 56% de corrélations attendues pour 88% chez le groupe contrôle (Oberman & Ramachandran, 2008). Cette diminution de l'effet est également présente chez les adultes dyslexiques, qui présentent des corrélations phonosymboliques significatives mais diminuées, soit 60% de corrélations correctes attendues pour 73% chez le groupe contrôle (Drijvers *et al.*, 2015).

Les auteures concernées suggèrent que ces effets sont liés à des difficultés dans le traitement multimodal. Il est donc particulièrement intéressant d'explorer cette piste : quel est le rôle de l'intégration multimodale dans le phonosymbolisme ? Où se placent les correspondances transmodales dans ce processus ?

1.3.3.2. Modélisation de l'intégration multimodale et des correspondances transmodales

Dans cette section, nous souhaitons interroger le lien entre les corrélations phonosymboliques, les correspondances transmodales et le mécanisme d'intégration multimodale. Pour ce faire, nous allons d'abord décrire le fonctionnement général de l'intégration multimodale selon le modèle d'Ernst (2005). Spence (2011) considère effectivement que le modèle Bayésien de l'intégration multimodale proposé par Ernst (2005) est l'outil de modélisation de l'intégration multimodale le plus adépte à expliquer l'implication des correspondances transmodales dans la perception générale. Ce modèle suggère que l'intégration multimodale résulte d'une mise en relation entre deux grandes estimations. Pour percevoir de la manière optimale, Ernst (2005) suggère que l'humain combine l'analyse :

- (1) d'indices effectifs dans le signal, tels qu'ils sont réellement perçus,
- (2) et de l'expérience statistique antérieure de l'environnement (le *coupling prior*).

Le processus d'intégration multimodale fonctionnerait donc sur la base d'un mécanisme prédictif : la personne percevante apprend que, d'une façon significative, certains signaux unimodaux sont corrélés et s'attend donc à ce que certains signaux apparaissent ensemble. Ces corrélations sont apprises par exposition systématique et sont une variable

statistique du modèle d'intégration multimodale proposée par Ernst, qui leur donne le nom de "*coupling prior*" (ou parfois simplement "*prior*", comme dans le modèle statistique bayésien classique). Lors de l'intégration de signaux, la robustesse du *coupling prior* peut induire des biais de perception. Si un *coupling prior* est très fort (significativement corrélé dans le contexte écologique de la personne), il modifiera la perception du signal réel par économie d'énergie. L'individu percevant aura moins utilité de traiter le signal réel, car la solution de l'intégration est plus rapide et précise. Cet effet est observable via un certain nombre d'illusions sensorielles bien connues, comme l'*effet McGurk* évoqué précédemment. Sur ce point, Ernst dit que "*la congruence transmodale peut induire des distorsions systématiques de l'expérience perceptuelle*" (2007 : 802). A l'inverse, si le *coupling prior* est faible ou que les signaux sont très clairs (sans bruit), alors la personne sera plus adepte à intégrer les indices effectifs du signal, et sera donc moins susceptible aux illusions sensorielles. En cela, le *coupling prior* répare les lacunes des signaux effectifs.

Pris en considération avec le reste de nos analyses, ces corrélations prédictives évoquent le phénomène des correspondances transmodales, et celles-ci (telles que aigu égal petit) pourraient être encodées dans le *coupling prior*.

Nous venons également de voir que la configuration du *coupling prior* est acquise par exposition à son environnement et serait donc un reflet de ce dernier. De plus, certaines associations sont fortes, d'autres sont faibles. Cela suggère que l'esprit encode robustement les associations permettant de percevoir le monde le plus optimalement possible. Ainsi, et si les correspondances transmodales sont effectivement liées au processus d'intégration, elles semblent universelles seulement sur la base du fait que l'humanité partage le même monde sensoriel et un mode de perception par défaut. Sur ce point, Spence (2011 : 988) dit que "*de telles correspondances [statistiques] ne seront évidemment pas universelles, car elles reflètent la réponse de l'organisme (par le biais de l'apprentissage associatif) aux statistiques idiosyncrasiques de l'environnement multisensoriel dans lequel il s'est récemment trouvé*"³.

Ce point de vue est justifié par le fait que certaines correspondances transmodales peuvent être apprises dans un contexte expérimental (Ernst, 2007). Les correspondances transmodales structurales, à priori à la source du phonosymbolisme, sont donc susceptibles de changer selon le champ perceptif de l'individu percevant. Lewkowicz (2014) théorise cette spécialisation multimodale sous le nom de "rétrécissement de la fenêtre perceptuelle multimodale", un phénomène développemental qui diminue la saillance perceptuelle de certaines catégories d'associations multisensorielles au profit d'autres. Cette spécialisation est une réponse adaptative à l'environnement de l'individu percevant : "*les effets régressifs de l'expérience précoce conduisent à l'ajustement du système perceptif du nourrisson humain, ce*

² Dans l'original : "cross-modal congruency can induce systematic distortions of perceptual experience" (Ernst, 2007:802)

³ Dans l'original : "However, that said, such [statistical] correspondences are clearly not going to be universal, in that they reflect the response of the organism (mediated by associative learning) to the idiosyncratic statistics of the multisensory environment in which it has recently found itself." (Spence, 2011:988)

qui aboutit finalement à une adéquation entre le système perceptif du nourrisson et les exigences de sa niche écologique quotidienne”⁴ (Lewkowicz, 2014 : 10).

1.3.4. Récapitulatif : influences à la fois innées et acquises

Pour reprendre nos hypothèses quant à l’origine cognitive du phonosymbolisme, nous avons vu que certaines analogies intersensorielles sont essentielles au fonctionnement de la perception en général et interviennent notamment dans le processus d’intégration multimodale et son mécanisme prédictif (Ernst, 2005). Parmi ces corrélations, il semblerait que certaines soient audiovisuelles : l’étude de Fort & Schwartz (2022) montre par exemple que nous faisons des équivalences intersensorielles entre des qualités de formes prototypiques et certaines qualités acoustiques. Ces analogies pourraient, elles-même, être motrices des corrélations phonosymboliques de forme.

Dans ce cadre, l’impression d’universalité de certaines corrélations (comme celles de forme) s’explique par le fait qu’elles émergent d’un mécanisme biologique inné commun aux humains neurotypiques – l’intégration multimodale – qui elle-même reflète indices sensoriels issus d’un monde physique que nous partageons tous. Les comportements divergents vis-à-vis des tendances phonosymboliques seraient donc le reflet du contexte écologique sensoriel des individus testés. Ce modèle, déployé à partir de l’articulation de diverses propositions de la littérature, est résumé dans la figure 4 ci-après. On ne peut néanmoins pas résumer le phonosymbolisme à ce modèle car, comme nous l’avons dit précédemment, d’autres éléments semblent se greffer à la robustesse de certaines associations, comme par exemple la graphie, qui semble renforcer l’*effet boubakiki* (Bottini *et al.*, 2018). Il est néanmoins intéressant de considérer ce modèle comme le socle phonosymbolique, auquel d’autres dimensions se greffent.

⁴ Dans l’original : “the regressive effects of early experience lead to the tuning of the human infant’s perceptual system and that this ultimately leads to a match between the infant’s perceptual system and the exigencies of its everyday ecological niche” (Lewkowicz, 2014:10)

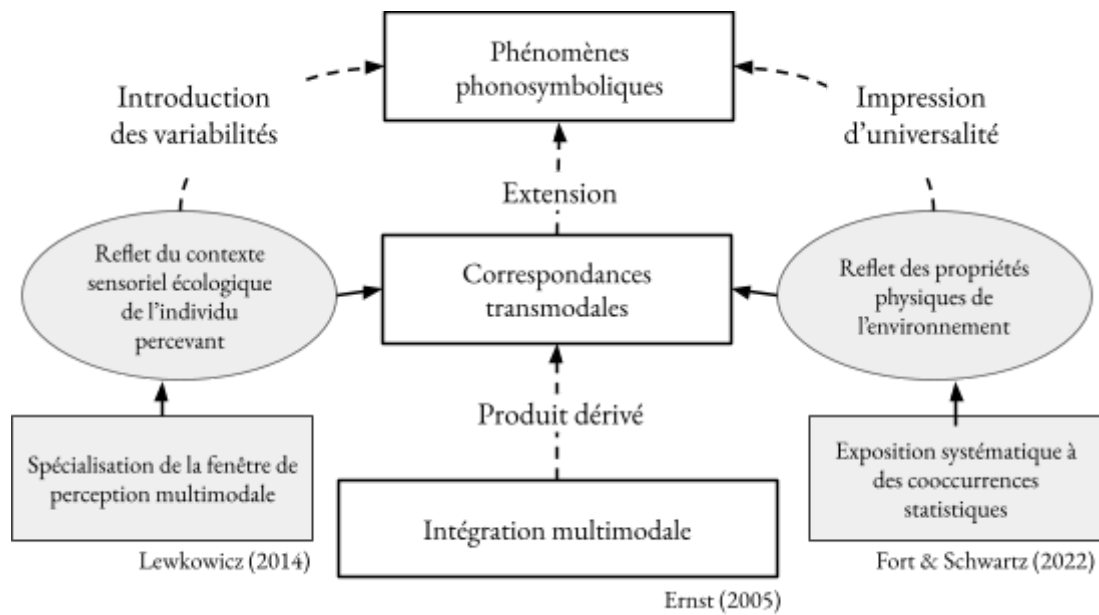


Figure 4. Schéma des mécanismes cognitifs hypothétiquement aux sources du phonosymbolisme.

Dans la suite de notre étude, nous souhaitons interroger ce modèle à travers l'étude des tendances phonosymboliques d'une population au contexte sensoriel atypique. Nous faisons ainsi le choix d'explorer les comportements phonosymboliques de populations déficientes visuelles pour explorer les questions de recherche suivantes :

- Quel est l'impact de la vue dans le phonosymbolisme de forme ?
- Les personnes non-voyantes expriment-elles des tendances phonosymboliques similaires à celles des personnes voyantes ?
- **In fine, des mondes sensoriels différents induisent-ils des comportements phonosymboliques différents ?**

2. Le phonosymbolisme chez les personnes déficientes visuelles

Dans cette partie, nous passerons en revue différentes littératures afin d'explorer trois axes principaux. D'abord, nous définirons les principaux enjeux sensoriels de la déficience visuelle. Ensuite, nous nous demanderons si les quatre dimensions intervenant dans la mise en place des corrélations phonosymboliques vues précédemment sont disponibles en situation de déficience visuelle. Ces fondements nous permettront de mieux aborder, en fin de partie, la littérature sur le phonosymbolisme chez les personnes non-voyantes.

2.1. Définitions et implications de la déficience visuelle

Dans cette première section, nous avons pour objectif de passer en revue les aspects essentiels de la perception chez les personnes non-voyantes. Le phonosymbolisme étant intrinsèquement sensoriel et multimodal, il semble primordial d'aborder ce problème en ayant une idée des enjeux sensoriels de la population étudiée.

Ainsi, nous aborderons la question des types développementaux de déficience visuelle ainsi que la réorganisation sensorielle unimodale et multimodale chez cette population, ayant toujours à l'idée de pouvoir réinvestir ces informations dans notre développement ultérieur.

2.1.1. Critères définitoires et types de déficiences visuelles

La déficience visuelle se définit très largement comme *“état d'une personne qui est privée de vision”*⁵. Cette condition est un continuum dont la variable fondamentale est l'acuité visuelle qui peut aussi bien être très faible que nulle, on parlera à ce moment de cécité. L'échelle de diagnostic de la déficience visuelle généralement citée est celle de l'Organisation Mondiale de la Santé, incluant cinq catégories comme suit :

	MALVOYANCE			CÉCITÉ	
	CATÉGORIE 1	CATÉGORIE 2	CATÉGORIE 3	CATÉGORIE 4	CATÉGORIE 5
Type déficience	MOYENNE	SÉVÈRE	PROFONDE	PRESQUE TOTALE	ABSOLUE
Acuité visuelle	$1/10 < x < 3/10$	$1/20 < x < 1/10$	$1/50 < x < 1/20$	$x < 1/50$	$x = 0/10$
			champ visuel inférieur à 10 degrés		absence totale de perception de lumière

Tableau 2. Échelle de la déficience visuelle selon l'OMS (2018). Les nombres représentent le degré de champ de vision du sujet.

⁵ CNRTL pour “cécité” consulté pour la dernière fois le 24/06/2025.

Car cette condition permet d'isoler la vue comme variable expérimentale, notre étude s'intéressera plutôt aux cas de cécité, soit l'absence de vision totale. Autrement, si la cécité est presque totale, le sujet ne perçoit que la lumière ou les mouvements d'ombres.

Nous ferons également la distinction entre cécité précoce ou tardive, un indice développemental de survenue de la déficience visuelle chez un individu. Dans le cas du phonosymbolisme, il a été démontré que l'âge d'*onset* de la déficience visuelle affecte les configurations phonosymboliques attestées, nous y reviendrons (Sourav *et al.*, 2019). Néanmoins, les définitions typologiques de la déficience visuelle précoce ou tardive sont hétérogènes dans la littérature. La cécité précoce inclut la cécité congénitale (naissance sans vision) et semble être définie comme survenant avant l'âge de 2 à 4 ans, alors que la déficience visuelle tardive survient plutôt après 3 à 5 ans (Bottini *et al.*, 2018; Deroy *et al.*, 2016; Fryer *et al.*, 2014; Hamilton-Fletcher *et al.*, 2018). Une étude neuroanatomique démontre que toute déficience visuelle acquise avant 16 ans induit des réorganisations corticales importantes (Sadato *et al.*, 2002), cependant, les réponses comportementales entre personnes à déficience visuelle précoce et tardive apparaissent déjà lorsque la vue est perdue après 12 ans (par ex. Burton *et al.*, 2002). Pour cette raison, nous nous fierons au critère de l'étude de Sourav *et al.* (2019) typologisant la déficience visuelle précoce comme la perte de la vue avant 12 ans et la déficience visuelle tardive comme ayant perdu la vue après 12 ans.

2.1.2. Réorganisation sensorielle et intégration multimodale chez les personnes non-voyantes

2.1.2.1. Paradigmes sensoriels dominants : implication dans l'analyse des formes

Dans les tâches phonosymboliques du type *bouba-kiki* ou *maluma-takete*, l'appréhension visuelle de l'aspect {rond/angulaire} d'une forme géométrique est essentielle à l'effet. En l'absence de vision, une personne non-voyante appréhende ces aspects géométriques par le toucher. Dans cette section, nous aborderons des considérations fondamentales pour distinguer la modalité visuelle de la modalité haptique.

Là où les personnes voyantes manipulent en majorité des images mentales liées à la vue, celles des personnes non-voyantes sont plutôt liées à la modalité haptique, d'ailleurs considérée comme "*la [modalité perceptive la] plus capable de supplanter la vision défaillante*" (Hatwell, 2003 : 12). La terminologie *haptique* a été proposée par Révész (1950) pour faire référence à la perception tactilo-kinesthésique, perception issue à la fois des sensations cutanées et des sensations proprioceptives. Hatwell (2003 : 121) explique que :

Les images mentales non visuelles des aveugles ont les mêmes propriétés générales fonctionnelles que celles des voyants. Elles sont cependant en partie spécifiques car le poids des perceptions auditives et tactiles dont elles dérivent y est évidemment plus important que chez les voyants. En conséquence, leur manipulation mentale est plus lente, plus laborieuse et parfois moins performante que celle des images à dominante visuelle des voyants, surtout lorsqu'il faut traiter une grande quantité d'informations spatiales ou des données typiquement visuelles comme les transformations dues à la perspective.

Malgré la difficulté que présente l'analyse spatiale pour une personne non-voyante, il a été montré que la perception haptique de formes 3D (Norman & Bartholomew, 2011) ou 2D (Heller *et al.*, 2003) est analogue chez les participants non-voyants précoces et voyants. En parallèle, les groupes non-voyants tardifs présentent de meilleures aptitudes dans cette tâche que les autres groupes. Il a été proposé que ceux-ci sont meilleurs dans ces tâches du fait d'une très bonne acuité haptique et d'images mentales visuelles antérieures, soit une forme de médiation visuelle de l'analyse haptique (Heller & Gentaz, 2018 : 21; Lederman *et al.*, 1990).

Il est cependant important de rappeler que, bien que l'haptique permette la perception géométrique, ce n'est pas ce qu'elle fait le mieux. La vue est bien plus performante dans ce cadre et permet une analyse globale des formes, de façon instantanée, peu coûteuse et dans un champ de perception étendu. A l'inverse, l'haptique permet une analyse séquentielle des caractéristiques d'une forme, via le contact physique direct dans un champ de perception restreint (Correa Silva *et al.*, 2014). La modalité haptique n'est donc pas spécialisée dans l'analyse géométrique et la condition d'exploration en contact haptique direct permet d'imaginer que les personnes non-voyantes sont moins confrontées aux formes géométriques que les voyants. Une mise en correspondance statistique de l'environnement entre des formes et des sons, envisagé ici comme mécanisme moteur du phonosymbolisme, n'existe peut-être pas à la même hauteur chez des personnes non-voyantes que chez les voyants.

Mascle (2020 : 36) écrit dans sa thèse que *“la vision serait spécialisée dans le traitement de données spatiales telle que la forme tandis que le système haptique serait spécialisé dans le traitement de propriétés matérielles comme la texture”*. Compte tenu de ces informations, il est intéressant d'envisager que les corrélations phonosymboliques de formes ne reflètent pas le contexte écologique des personnes non-voyantes de façon analogue à une personne voyante, du fait d'une divergence de paradigmes sensoriels dominants.

2.1.2.2. Réorganisation sensorielle : l'audio-haptique en l'absence de vision

Comme évoqué précédemment, il est aujourd'hui accepté que la privation visuelle induit des réorganisations corticales plus ou moins importantes selon l'âge de perte de vue. Par exemple, la cécité précoce entraîne le réinvestissement du cortex visuel pour des activités

haptiques, notamment la lecture du Braille (Sadato *et al.*, 1996). Néanmoins, notre étude n'a pas l'ambition d'explorer les aspects neuro-anatomiques de la déficience visuelle ou du phonosymbolisme. Dans cette section, nous aborderons rapidement la réorganisation sensorielle observable comportementalement chez les personnes non-voyantes. Scheller *et al.* (2021) résumant dans leur étude deux conséquences sensorielles importantes de la cécité.

D'abord, l'absence de vue induit des compensations transmodales. Cet aspect plutôt positif est bien étudié pour l'audition et l'haptique, bien que les avantages démontrés pour le toucher soient plutôt controversés et certainement liés à l'entraînement tactile (Sathian & Stilla, 2010). Cela étant dit, les compensations auditives sont certaines. Par exemple, les personnes non-voyantes discriminent des énoncés oraux à vitesse d'énonciation autrement incompréhensible chez un sujet typique (Moos *et al.*, 2008) et sont plus performantes dans des tâches de discriminations des fréquences (Gougoux *et al.*, 2004). Elles discriminent également plus rapidement les voyelles (Ménard *et al.*, 2009), notamment celles labialisées (Hirsch *et al.*, 2010 : 5), ce probablement car elles *“développeraient une sensibilité auditive plus prononcée leur permettant de percevoir des changements plus fins au niveau de la limite inférieure du bruit de friction”*.

Généralement, les personnes non-voyantes ont des temps de réaction plus courts que des voyants dans des tâches de décision lexicale, d'en moyenne 80 millisecondes (Schild & Friedrich, 2018). L'apparition d'une N400 lors de la perception d'une phrase incongruente en EEG est également plus précoce chez les personnes non-voyantes que les voyants, laissant entendre que le traitement général de la phrase est plus rapide pour eux (Röder *et al.*, 2000). Les personnes non-voyantes ont également une meilleure capacité à percevoir les objets et repères spatiaux via à la réverbération du son que les voyants (Arias, 1996).

Nonobstant, l'absence de vue induit des réorganisations plutôt négatives que Scheller *et al.* (2021) attribuent à la calibration intersensorielle, une hypothèse développementale de la perception supposant que certains sens dominent les autres et sont donc responsables de les calibrer afin qu'ils perçoivent optimalement lors de tâches spécifiques. Par exemple, la vue est le sens le plus précis dans l'analyse spatiale et permettrait donc aux autres sens d'obtenir les informations nécessaires à une perception optimale. Gori *et al.* (2010) montrent par exemple que l'absence de vue induit une discrimination moindre de l'orientation d'objets. Selon eux, l'absence de vision aurait donc des répercussions sur la perception spatiale.

2.1.2.3. Intégration multimodale en l'absence de vision

Nous venons de voir qu'il y a réorganisation sensorielle chez les personnes déficientes visuelles. Celle-ci affecte également des mécanismes perceptifs multimodaux, notamment l'intégration multimodale. Dans cette section, nous explorons ces réorganisations en interrogeant de la susceptibilité des personnes non-voyantes aux illusions sensorielles audio-haptiques, à priori provoquées par la prévalence du mécanisme prédictif de l'intégration multimodale (Ernst, 2005).

Deux études montrent que les personnes non-voyantes précoces sont moins sensibles aux illusions sensorielles que des groupes non-voyants tardifs ou voyants. C'est par exemple le cas pour des illusions liées au repérage spatial comme l'effet ventriloque (Occelli *et al.*, 2012) ou des tâches de discriminations audio-haptiques plus générales (Hötting & Röder, 2004). Ces études suggèrent que les personnes non-voyantes précoces sont moins susceptibles à l'interférence multimodale car la compensation sensorielle perfectionne le traitement unimodal au stade de ne pas avoir de bénéfice dans l'usage du mécanisme prédictif. Autrement dit, cette population aurait tendance à traiter les signaux de façon unimodale (Hötting & Röder, 2004), tendance déjà observée chez les animaux aveugles (Wallace, 2004).

A l'inverse, une étude plus récente mesurant l'effet du son sur la perception de la taille d'un objet montre que les non-voyants précoces sont aussi sensibles à la tâche que des personnes voyantes. Ici, ce sont les personnes déficientes visuelles tardives qui sont peu susceptibles à l'illusion (Scheller *et al.*, 2021). Enfin, une quatrième étude a montré que les individus non-voyants précoces et tardifs sont moins sensibles à l'effet *Parchment-Skin* que des voyants, une illusion audio-haptique mesurant l'effet du bruissement sur la perception de la texture cutanée (Champoux *et al.*, 2011)

Ainsi, nous pouvons voir que les résultats de cette littérature sont plutôt hétérogènes et témoignent à minima d'un traitement multimodal atypique chez les personnes non-voyantes, dépendant du domaine testé, de la tâche et de l'*onset* de privation visuelle. Ces observations ne sont pas suffisantes pour tirer des conclusions quant à l'impact de cette atypicité sur notre recherche. A ce stade, nous pouvons conclure que le mécanisme semble fonctionnel pour les deux groupes, mais ne provoque pas systématiquement des illusions (moins de coactivation intersensorielle dans certains contextes).

2.2. Implication de la vue dans le phonosymbolisme

Dans la section précédente, nous avons vu que la cognition sensorielle des personnes non-voyantes se retrouve largement modifiée par l'absence de vision. Pour revenir à notre problématique initiale, nous nous demanderons maintenant si les personnes déficientes visuelles ont accès aux différentes dimensions influençant les configurations des corrélations phonosymboliques chez un sujet typique malgré cette réorganisation sensorielle. Nous suivrons le modèle proposé dans la première partie, soit exploration de la dimension articulatoire, acoustique, orthographique et de l'expérience sensorielle écologique.

Dans un premier temps, la dimension articulatoire étant souvent considérée essentielle aux analogies phonosymboliques, il convient de se demander si les personnes déficientes visuelles ont une phonation analogue à celle d'un sujet typique. Plusieurs études suggèrent effectivement que la vue oriente et calibre le bon apprentissage des gestes articulatoires en fournissant des indices quant au mouvement des lèvres et de la mâchoire lors de l'articulation (voir par exemple l'*effet McGurk* ou l'importance des indices visuels lors de la perception de la parole dans le bruit, p. ex. Sumbly & Pollack, 1954). Confirmant ces hypothèses, il a été montré que les personnes non-voyantes ont une implémentation divergente des gestes articulatoires typiques du français, notamment moins de labialisation, compensée par une rétraction plus importante de la langue (Ménard *et al.*, 2013). De pareilles études sur la production des consonnes n'existent pas à notre connaissance, ce qui est dommage pour notre étude compte tenu du poids phonosymbolique des consonnes (Nielsen & Rendall, 2011).

Pour faire le pont avec la dimension acoustique, Ménard *et al.* (2013) remarquent également que les écarts inter-formantiques moyens des voyelles produites par des voyants sont plus importants que ceux des personnes non-voyantes. Cela suggère que, bien que les personnes déficientes visuelles discriminent mieux les voyelles que les voyants sur le plan perceptif (Ménard *et al.*, 2009), ils présentent moins de contraste acoustique dans leur production. Quant à la perception du signal sonore en général, nous avons déjà vu que les personnes non-voyantes sont très sensibles aux signaux sonores et discriminent mieux les jeux de fréquences que leurs pairs voyants (Gougoux *et al.*, 2004).

Pour continuer, interrogeons la dimension orthographique. Nous avons vu plus tôt que l'orthographe joue une part importante, même si elle n'est pas cruciale, dans la configuration des corrélations phonosymboliques. Comme il a été soulevé par Bottini *et al.* (2018), les personnes non-voyantes ont eu un accès réduit ou absent à l'écriture visuelle (aussi dit lire "*en noir*") selon l'*onset* de leur déficience visuelle. Néanmoins, les personnes non-voyantes ont une orthographe adaptée à la modalité haptique, le braille, qu'il serait dommage d'ignorer. Le braille est un système d'écriture où une lettre correspond à une combinaison de points en reliefs inscrits dans une matrice rectangulaire de deux points sur trois. La lecture

braille se fait par balayage tactile des points et il semblerait que l'unité de perception du braille soit des *“patterns de cisaillements latéraux de la densité des points”* (Millar, 1997 : 172). Quant à la réalité cognitive des lettres brailles en mémoire, il semblerait que, *“tout comme le nom et le son des lettres, les caractéristiques physiques et tactiles de la “texture” sont codées dans la mémoire”* (Millar, 1997 : 199). Les corrélations phonosymboliques potentielles des personnes non-voyantes pourraient donc être affectées par des analogies orthographiques, avec changement de paradigme sensoriel de la vision à la texture.

Enfin, nous avons suggéré que le mécanisme moteur du phonosymbolisme est lié aux cooccurrences statistiques disponibles dans l'environnement. Or, et cela est soutenu par toutes les considérations énoncées dans les sections précédentes, il y a réorganisation des indices sensoriels pertinents dans le contexte écologique des personnes déficientes visuelles. Cela affecte notamment les représentations mentales, qui sont majoritairement liées à la vue chez les voyants et à l'haptique chez les déficients visuels (Hatwell, 2003), mais également les cooccurrences statistiques disponibles dans l'environnement. Il est tout de même intéressant de considérer que l'expérience sensorielle directe n'est pas nécessaire à la création d'associations, par exemple, les *“cooccurrences statistiques dans la parole permettent de saisir les associations sémantiques entre les couleurs, indépendamment de l'accès à la vision”*⁶ (Campbell & Bergelson, 2022 : 4).

Au terme de cette section, nous avons pu constater que l'absence de vision affecte la plupart des catégories entendues comme essentielles à la mise en place des corrélations phonosymboliques. A partir des informations énoncées dans cette partie, il est intéressant de se demander la mesure dans laquelle ces changements affectent réellement les tendances phonosymboliques des personnes déficientes visuelles. C'est cette question qui nous intéressera dans la partie suivante.

2.3. Réponses comportementales de la population non-voyante aux correspondances transmodales

Dans cette dernière partie, nous nous demanderons dans quelle mesure l'absence de vue affecte l'émergence de correspondances transmodales non-linguistiques et, à terme, l'émergence des tendances phonosymboliques chez les populations non-voyantes.

⁶ Dans l'original : *“spoken language co-occurrence statistics capture color-semantic associations regardless of access to vision”*

2.3.1. Correspondances transmodales simples chez les personnes déficientes visuelles

Nous avons considéré précédemment que l'émergence du phonosymbolisme s'appuie sur des correspondances transmodales non-linguistiques. Dans cette partie, nous ferons un état de l'art de l'expression des ces associations, impliquant généralement des jeux de fréquence en tant que stimuli sonores, chez les personnes déficientes visuelles. Sont-elles analogues à celles des voyants ? Une revue de la littérature permet de dégager quatre grandes tendances.

D'abord, les personnes déficientes visuelles semblent parfois exprimer les mêmes associations que des personnes voyantes. C'est le cas pour les correspondances transmodales entre une fréquence et la taille d'un objet, où tous les groupes expriment les corrélations {aiguë~petit} et {grave~grand}. C'est également le cas pour les corrélations entre une fréquence et un poids, où {aiguë~léger} et {grave~lourd} (Hamilton-Fletcher *et al.*, 2018).

Il arrivé également que les personnes non-voyantes n'expriment pas les correspondances observées chez des personnes voyantes. C'est par exemple le cas pour les associations entre une fréquence et la forme d'un objet, où les personnes non-voyantes précoces ne semblent pas faire la corrélation {aiguë~angulaire} ou {grave~rond} (Hamilton-Fletcher *et al.*, 2018).

De façon intéressante, il arrive également le cas de figure opposé, où les personnes non-voyantes expriment des correspondances absentes chez les personnes voyantes. Par exemple, elles semblent être les seules à exprimer une corrélation significative entre une fréquence et une texture, où {grave~lisse} et {aiguë~rugueux} (Hamilton-Fletcher *et al.*, 2018). Un récapitulatif des données de cette étude est donnée en tableau 3.

	fréquence~taille	fréquence~poids	fréquence~forme	fréquence~texture
	{grave} ~ {aigu} {grand}~{petit}	{grave} ~ {aigu} {lourd}~{léger}	{grave} ~ {aigu} {rond}~{angulaire}	{grave} ~ {aigu} {rugueux}~{lisse}
Non-voyant	effet significatif	effet significatif	effet non significatif	effet significatif
Voyant	effet significatif	effet significatif	effet significatif	effet non significatif

Tableau 3. Résultats de Hamilton-Fletcher et al. (2018), corrélations transmodales évaluées entre des sons purs et d'autres qualités perceptuelles, par groupe.

Enfin, il arrive que pour une même opposition de traits d'une correspondance, les paradigmes changent entre le groupe voyant et le groupe déficient visuel. Par exemple, dans la relation entre indice acoustique et perception de la distance sur le plan horizontal, les voyants associent l'élévation de la fréquence moyenne d'un son pur avec l'augmentation de la distance avec un objet (sur le plan horizontal). Or, chez les déficients visuels, plus la fréquence augmente, plus l'objet s'approche (Eitan *et al.*, 2012).

En conclusion, il semblerait que les individus non-voyants expriment des correspondances transmodales, bien qu'elles ne soient pas toujours celles exprimées par les voyants et inversement. Clairement, la vue semble médier la configuration (voire l'émergence) de certaines correspondances chez les voyants. Il est néanmoins très intéressant de remarquer que l'absence de vision semble également avoir cet effet, en permettant l'émergence de correspondances transmodales de texture.

La réorganisation des paradigmes de correspondance transmodales non-linguistiques entre les groupes est particulièrement intéressante à considérer, si ces correspondances sont réellement à la source du phonosymbolisme. Dans la suite de ce papier, nous ferons un état de l'art des connaissances concernant les corrélations phonosymboliques des personnes déficientes visuelles.

2.3.2. Effet bouba-kiki absent chez les personnes déficientes visuelles

2.1.2.3. Effet absent

Fryer *et al.* (2014) sont les premiers à avoir testé la corrélation phonosymbolique de forme chez des personnes non-voyantes, en recréant le protocole de Köhler (1929) en modalité haptique. La moitié de leurs participants étaient voyants et l'autre moitié déficients visuels de différents types (cécité congénitale, $n = 6$; cécité tardive, $n = 17$; malvoyance, $n = 19$). L'effet que nous allons décrire a été observé même lorsque le groupe déficient visuel est mieux équilibré, comme chez Piller *et al.* (2023) où les 17 participants sont non-voyants précoces. Ces chercheur·es ont demandé à leurs participants de plonger la main dans un sac contenant deux formes 3D présentant des caractéristiques {rondes} ou {angulaires} inconnues (voir figure 5) et d'extraire la forme qui correspond le mieux à /buba/ ou /kiki/.

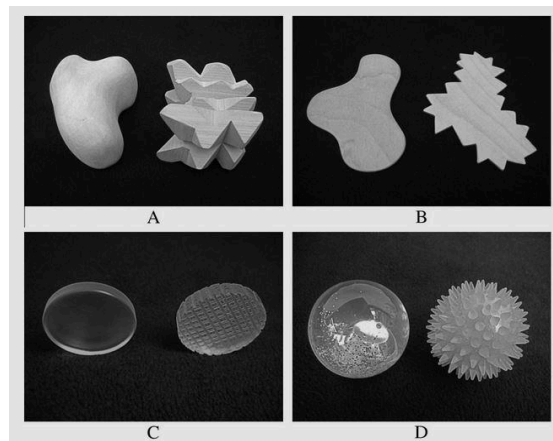


Figure 5. Objets de forme utilisés dans la tâche de Fryer et al. (2014)

Leur étude montre que les participants voyants font la corrélation attendue (forme {ronde} pour /buba/ et inversement) de façon significative, soit 84% de réussite chez Fryer *et al.* (2014). Néanmoins, cet effet ne s'observe pas chez les participants déficients visuels qui sélectionnent les formes à hauteur de chance (50% de réussite pour cécité précoce, 55% pour la cécité tardive). Ainsi, l'*effet bouba-kiki* semble absent chez les personnes déficientes visuelles précoces et très faible chez celles tardives. Cette absence paraît néanmoins liée à une période critique d'accès à la vision dans l'enfance. Une étude de Sourav *et al.* (2019) montre effectivement que l'absence d'*effet bouba-kiki* n'émerge que si la cécité se signale avant l'âge de 12 ans. Les personnes ayant perdu la vue après cet âge présentent quant à elles des tendances phonosymboliques analogues à des voyants.

Ces observations, prises avec les résultats de Hamilton-Fletcher *et al.* (2018) démontrant que les personnes déficientes visuelles précoces ne font pas de corrélation systématique entre une forme et une variation de hauteur acoustique, suggèrent que l'expérience visuelle permet en partie l'émergence de corrélations phonosymboliques de forme. Du moins, l'absence de la vue affaiblie de façon très conséquente l'effet autrement observable de façon robuste chez des participants voyants.

2.1.2.3. Effet latent

Le comportement phonosymbolique atypique des participants non-voyants précoces ne se limite pas à la précision de réussite globale. Graven & Desebrock (2018) observent également un effet d'amélioration en cours de tâche pour les participants non-voyants. Ces chercheur-es ont demandé à un groupe déficient visuel précoce et voyant (travaillant sans voir) d'appareiller des contours de forme 2D aux stimuli /buba/, /kiki/, et des sons purs imitant les caractéristiques de hauteur et de durée de ces logatomes. Cette passation s'est déroulée en huit tours. Les taux de réponses correctes au niveau phonosymbolique au premier tour sont similaires aux résultats présentés par Fryer *et al.* (2014), c'est-à-dire : 50% de réussite chez le groupe déficient visuel et 75% chez les voyants. Néanmoins, d'ici le dernier tour, ces premiers résultats s'inversent : 75% de réussite chez le groupe déficient visuel et 50% chez les voyants. C'est à partir du quatrième trial que cette tendance se développe. Or, l'étude de Fryer *et al.* (2014) et des chercheurs ayant répliqué leur étude (Sourav *et al.*, 2019; Piller *et al.*, 2023) se sont généralement arrêtés au bout de un à quatre trials.

Ces résultats sont difficiles à interpréter. Ils suggèrent dans un premier temps que, bien qu'il n'y ait pas d'effet bouba-kiki instantané chez les personnes non-voyantes, cet effet émerge tout de même avec le temps et serait donc en quelque sorte appris. A l'inverse, pour les voyants travaillant sans voir, l'effet est instantané mais ne perdure pas dans le temps.

Graven & Desebrock (2018) suggèrent que cet effet est lié au mode de traitement haptique utilisé entre les participants voyants et déficients visuels. Graven & Desebrock (2018 : 203) suggèrent que les voyants arrivent à analyser haptiquement les propriétés pertinentes des formes grâce à leurs compétences visuelles, mais qu'ils ne les invoquent pas à chaque tour, "*comme s'ils étaient capables de faire appel à leurs représentations mentales visuelles de manière exceptionnelle/isolée*"⁷, considérant que la tendance observée résulte du manque d'expérience dans l'exploration haptique chez les voyants. Quant à la tendance d'apprentissage progressif des participants non-voyants, ils posent l'hypothèse que la qualité de traitement sonore étant nettement plus rapide que le traitement haptique d'une forme, il a nécessité plus de temps au groupe non-voyant pour comprendre l'opposition sur laquelle ils étaient testés.

2.3.3. Effet maluma-takete présent mais réduit chez les personnes déficientes visuelles

Pour continuer, il est important de souligner que les études remarquant une absence de phonosymbolisme de forme chez les personnes non-voyantes (p. ex. Fryer *et al.*, 2014; Sourav *et al.*, 2019; Piller *et al.*, 2023) observent précisément une absence d'*effet bouba-kiki*. Leurs résultats sont donc éclairants mais restrictifs, car ils n'interrogent que le ressenti phonosymbolique de deux logatomes.

Afin d'examiner les ressentis phonosymboliques hors de ce paradigme, Bottini *et al.* (2018) ont reproduit le protocole de Fryer *et al.* (2014) avec les logatomes /maluma/ et /takete/, plutôt que /buba/ et /kiki/. Leurs résultats montrent que, dans cette condition, les personnes non-voyantes congénitales présentent une corrélation phonosymbolique significative (~73% de réussite), bien qu'un peu moins systématique que les participants voyants (~83% de réussite). Les stimuli proposés par Ramachandran & Hubbard (2001) ne permettent donc pas une opposition phonosymbolique suffisante chez les personnes non-voyantes pour polariser leurs sentiments de façon significative. Selon Bottini *et al.* (2018), l'opposition consonantique entre /b/ et /k/ n'est pas phonosymboliquement pertinente pour les groupes non-voyants (alors qu'elle l'est pour les voyants) car, pour eux, ces deux consonnes évoquent la même qualité de forme {angulaire}.

Dans la suite de leur étude, Bottini *et al.* (2018) posent l'hypothèse que les facteurs d'iconicité chez les individus voyants et non-voyants sont réorganisés, expliquant l'absence d'*effet bouba-kiki*. Ces chercheurs mettent alors en place une tâche d'association en choix contraint avec 240 pseudo-mots contrastés sur le mode articulaire et le voisement des consonnes (en plus de la labialisation des voyelles) afin d'obtenir un coefficient d'iconicité

⁷ Dans l'original : "*as if they were able to call upon their visual imagery as a 'one off'*"

de différentes classes articulatoires. La figure ci-dessous illustre la force des sentiments phonosymboliques des classes consonantiques ou vocaliques étudiées par Bottini *et al.* (2018). Lorsque la barre dépasse la ligne rouge, le groupe phonémique est perçu {angulaire}.

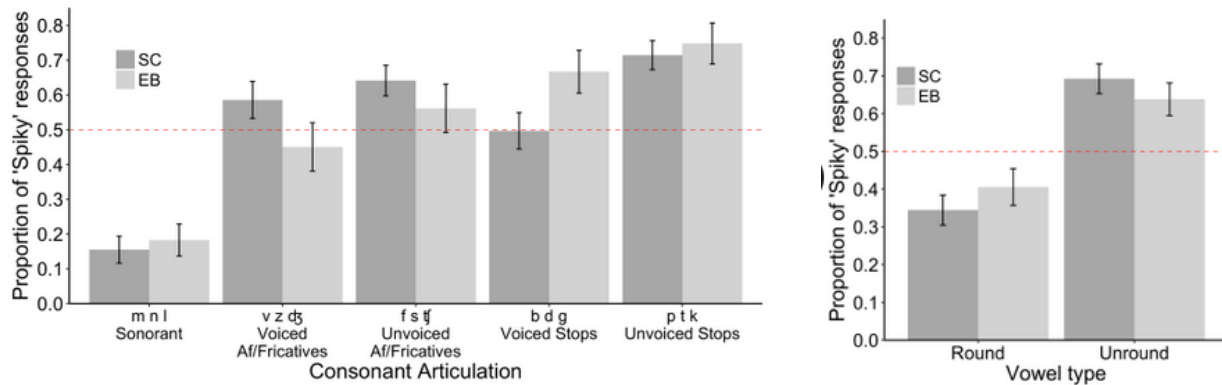


Figure 6. Résultats de Bottini et al. (2018) comparant les corrélations phonosymboliques d'iconicité angulaire chez les voyants (groupe SC) et les non-voyants précoces (groupe EB).

Dans cette figure, on remarque que l'association des consonnes avec les qualités d'angularité croît d'une façon linéaire selon le critère de [continuité], puis [sonorité] pour le groupe non-voyant. Bottini *et al.* (2018) suggèrent que ce n'est pas le cas pour les voyants, car leurs sentiments sont affectés par leurs représentations orthographiques. De plus, la [continuité] semble prévaloir sur la sonorité : les consonnes les plus {angulaires} sont obstruantes. Néanmoins, une consonne obstruante [-voisée] est plus {angulaire} que son pendant [+voisée]. De même, la nasalité semble être le caractère articulatoire le plus {rond}. Quant aux résultats de Fryer *et al.* (2014), ils sont expliqués par les oppositions phonosymboliques des consonnes déterminantes dans l'effet *bouba-kiki* : on remarque bien que /b/ et /k/ partagent le même trait [-continu], ce qui malgré les oppositions de lieu d'articulation et de voisement, donne un caractère à dominante {angulaire} à /buba/ et /kiki/ pour les personnes non-voyantes.

Pour ce qui est des voyelles, on remarque que le groupe non-voyant associe de façon moins significative la labialisation avec le trait {rond} que les voyants. Étant donné qu'ils ne produisent pas autant d'arrondissement des lèvres lors de la phonation que les voyants (Ménard *et al.*, 2013), et que le symbolisme phonétique semble en partie basé sur des analogies intersensorielles d'ordre kinésique, il est possible que l'association soit moins significative au niveau phonosymbolique car le geste articulatoire est moins bien perçu. Cette observation va dans le sens de l'hypothèse labiale de Ramachandran & Hubbard (2001), qui avancent qu'il y aurait une contiguïté entre les réseaux neuronaux traitant l'information visuelle des contours des formes et des contours des lèvres lors de l'articulation, résultant en la corrélation du trait sémantique {rond} avec les phonèmes [+arrondi].

Contrairement aux études précédentes, l'étude de Bottini *et al.* (2018) montre bien que les corrélations phonosymboliques ne sont pas absentes chez les personnes non-voyantes, mais qu'elles répondent en fait à des facteurs d'iconicité différents de ceux des voyants. En effet, contrairement à l'*effet bouba-kiki*, l'*effet maluma-takete* semble réduit mais existant chez les personnes non-voyantes. Sur ce point, il semblerait qu'une expérience différente de l'environnement affecte les représentations phonosymboliques.

Ces remarques, prises avec les observations sur correspondances transmodales non-linguistiques évoquées précédemment, sont particulièrement intéressantes et nous permettent d'envisager plusieurs hypothèses quant à la suite de cette étude.

3. Perspectives, question de recherche

Nous venons de voir que les personnes non-voyantes présentent un profil atypique en matière de perception phonosymbolique. Les résultats autour de cette littérature sont hétérogènes, observant à la fois une absence de sentiment phonosymbolique (p ex. Fryer *et al.*, 2014) et une réorganisation des corrélations phonosymboliques autrement observables chez une population typique (Bottini *et al.*, 2018). Ces données comportementales montrent à minima que la réorganisation sensorielle chez les personnes non-voyantes influence la perception des effets phonosymboliques. Notre étude vise à explorer l'effet de cette réorganisation sensorielle, en nous focalisant sur les questions de recherche suivantes :

- **La vue est-elle nécessaire à l'émergence du phonosymbolisme ?**
- **Est-ce que des corps différents induisent des représentations phonosymboliques différentes, et si oui, dans quelle mesure ?**
- **Pourquoi les personnes non-voyantes n'expriment-elles pas les corrélations phonosymboliques de façon analogue à un voyant ?**

Plusieurs hypothèses émergent de la littérature. Nous tâcherons de les formuler à partir des informations évoquées dans les parties précédentes.

3.1. Hypothèse de la réorganisation des mécanismes phonosymboliques

Dans un premier temps, il a été démontré que le public déficient visuel exprime des correspondances transmodales parfois différentes des personnes voyantes (Hamilton-Fletcher

et al., 2018) pointant vers une réorganisation des paradigmes phonosymboliques chez les cette population, notamment observée en condition de forme (Bottini *et al.*, 2018). Ici, nous proposons que cette divergence est directement liée à la réorganisation sensorielle.

En effet, l'input sensoriel pertinent des personnes voyantes est généralement lié à la vue, alors que l'haptique est dominante chez les personnes non-voyantes. De plus, nous avons vu que ces modalités ont des domaines de spécialité différents (forme/spatialité vs texture/matérialité), laissant entendre que chaque groupe se repose sur des indices pertinents très différents pour percevoir optimalement leur environnement. Ces niches écologiques divergentes impliquent une adaptation spécifique des mécanismes perceptifs de chaque groupe. Notamment, le fonctionnement du mécanisme prédictif d'intégration multimodale doit être différent, car il se développe par exposition à son propre environnement (Ernst, 2007). Ainsi, nous supposons que les correspondances transmodales pertinentes à l'appréhension du monde sont atypiques chez les personnes non-voyantes du fait de leur niche écologique, induisant par extension des tendances phonosymboliques atypiques.

Si l'atypicité du groupe non-voyant est strictement liée à l'input sensoriel, il est alors intéressant d'envisager cette réorganisation phonosymbolique en termes de la répartition de la force des *coupling priors*. Dans ce cas, la différence entre le groupe voyant et non-voyant serait exclusivement lié à la mise en place des correspondances phonosymboliques de forme. Les études abordées précédemment suggèrent effectivement qu'il n'y a pas beaucoup de paramètres à l'œuvre dans la mise en place du phonosymbolisme de forme chez les personnes déficientes visuelles. Comme nous l'avons vu plus tôt, il a été montré que les personnes non-voyantes ne sont pas sensibles aux correspondances transmodales entre des sons purs et des formes, alors que les voyants le sont. Cela suggère que le *couling prior* entre ces deux percepts est faible. Pourtant, Bottini *et al.* (2018) montrent que les personnes non-voyantes expriment des corrélations phonosymboliques de forme de façon significative, suggérant que le seul codage analogique entre un son et une forme chez cette population se fait via les caractéristiques articulatoires de phonèmes. Nous suggérons notamment que l'iconicité phonosymbolique de forme est moins forte chez les personnes déficientes visuelles car elles sont peu confrontées à des associations son/forme au quotidien. Nonobstant, un patrimoine perceptuel de cooccurrences statistiques entre des sons et d'autres percepts doit bien exister pour les personnes non-voyantes à une autre échelle.

De façon intéressante pour nous, les personnes non-voyantes sont sensibles à des correspondances entre des sons purs et des textures, alors que les voyants ne le sont pas (Hamilton-Fletcher *et al.*, 2018). Nous posons ainsi l'hypothèse suivante, illustrée par la figure 7 : une variation d'input sensoriel entre les personnes voyantes et non-voyantes induit une variation de l'output phonosymbolique. Si les cooccurrences statistiques sont à la source du phonosymbolisme, nous suggérons que les personnes non-voyantes ont de bonnes représentations phonosymboliques de texture, de façon analogue voire supérieure aux voyants.

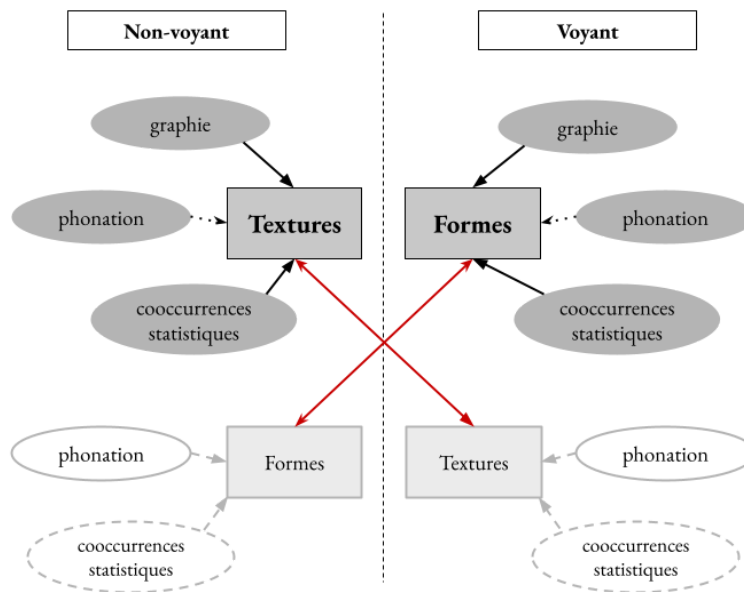


Figure 7. Dimensions intervenant dans le degré d'iconicité phonosymbolique de forme et de texture chez les personnes non-voyantes ou voyantes.

3.2. Hypothèse du déficit d'intégration multimodale

Dans un second temps, il a été montré que, bien que les correspondances phonosymboliques puissent s'exprimer dans des contextes particuliers, les effets phonosymboliques sont toujours plus faibles chez les participants non-voyants que pour les participants voyants. Pris avec le fait que l'*effet bouba-kiki* n'émerge pas chez les personnes atteintes de cécité (Fryer *et al.*, 2014), il a fréquemment été suggéré que l'absence de vue induit une absence de tendances phonosymboliques en général.

De façon intéressante, les groupes non-voyants ont fréquemment un comportement similaire à ceux observés chez des personnes dyslexiques ou sur le spectre autistique, c'est-à-dire un *effet bouba-kiki* réduit ou absent. Or, ces deux troubles sont marqués par une atypicité dans le traitement multimodal. En parallèle, nous avons vu que les personnes non-voyantes présentent un mécanisme d'intégration multimodale atypique, laissant supposer que les personnes déficientes visuelles n'expriment pas de corrélations phonosymboliques type bouba-kiki car l'absence de vue induit une atypicité dans le mécanisme d'intégration multimodale, entravant l'émergence de corrélations phonosymboliques chez cette population. Selon cette hypothèse l'intégration multimodale est affectée négativement par l'absence de vision et les individus auront tendance à traiter les signaux de façon unimodale plutôt que multimodale (p ex. Hötting & Röder, 2004).

Nous posons ainsi notre seconde hypothèse : l'absence de vision affecterait le fonctionnement typique de la perception multimodale, qui par cascade induit des correspondances transmodales faibles et une réduction de la sensibilité aux phénomènes phonosymboliques en général chez les personnes non-voyantes.

3.3. Visualisation des hypothèses

Pour reprendre les informations énoncées ici, nous venons donc à deux grandes hypothèses quant à la question : pourquoi les personnes déficientes visuelles n'expriment-elles pas les corrélations phonosymboliques de façon analogue à un voyant ?

Dans un premier temps, la divergence entre les groupes se situerait plutôt au niveau de l'input, où une variation de l'input sensoriel contribuant à la mise en place acquise du phonosymbolisme induirait des comportements de sortie divergents selon les groupes. Cette hypothèse est illustrée dans la figure ci-dessous en (1).

Dans un second temps, la divergence entre les groupes se situerait plutôt au niveau d'une atypicité du mécanisme aux racines du phonosymbolisme, l'intégration multimodale. Cette atypicité causerait un déficit généralisé de tendances phonosymboliques chez les personnes déficientes visuelles. Cette hypothèse est illustrée dans la figure ci-dessous en (2).

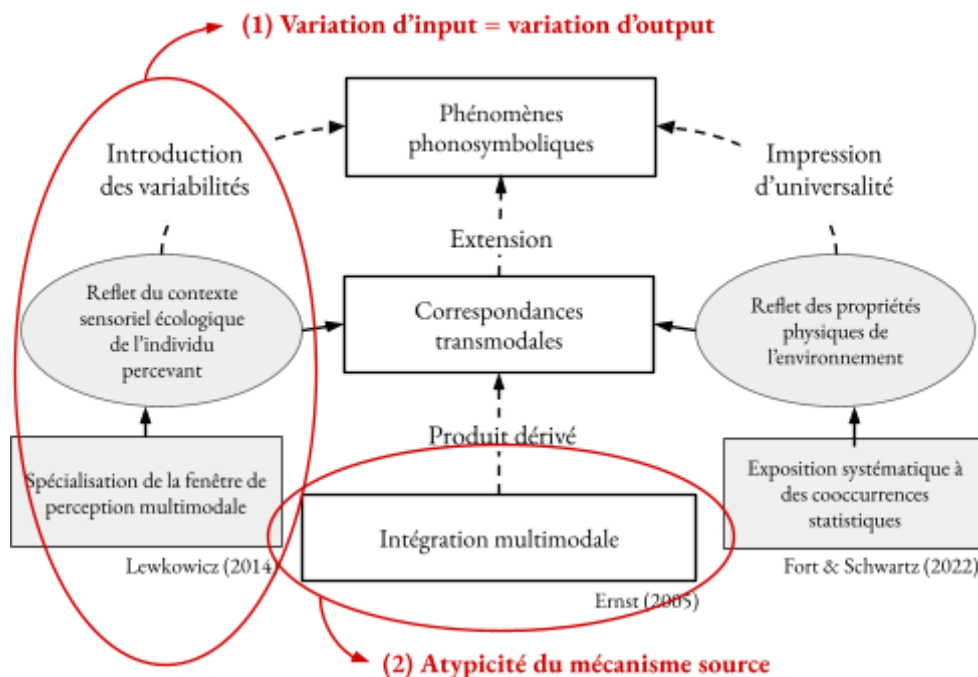


Figure 8. Visualisation des zones hypothétiques de divergence comportementale entre le groupe de personnes voyantes et non-voyantes, appliquées au modèle des mécanismes cognitifs aux sources du phonosymbolisme.

Chapitre 2 : Protocole expérimental

Comme nous l'avons vu au cours du chapitre précédent, la différence de comportement phonosymbolique entre des groupes d'individus voyants et déficients visuels a déjà été démontrée dans la littérature via des tests expérimentaux dits *explicits*, soit se reposant sur une manipulation consciente du participant dans des tâches d'association logatome/objet en choix contraint. Ici, nous avons testé si cette différence comportementale s'étend à un test mesurant la force des corrélations phonosymboliques de participants de façon *implicite*, soit en mesurant des données dérivées de l'avantage cognitif que présentent les items linguistiques phonosymboliques.

Pour ce faire, une tâche de mémorisation de paires logatome/objet congruentes ou incongruentes sur critères phonosymboliques a été mise en place. Dans ce chapitre, nous décrivons le projet expérimental voué à l'exploration des hypothèses établies précédemment.

1. Choix de la tâche : reconnaissance de paires

Dans cette partie, nous explicitons les fondements théoriques ayant motivé le choix de la tâche expérimentale.

1.1. Généralités sur la tâche de reconnaissance de paires

Une tâche de reconnaissance d'associations consiste généralement en deux phases : une phase d'apprentissage (encodage) d'associations nouvelles, suivie d'une phase de reconnaissance des paires apprises. Dans cette deuxième phase, on présente des paires intactes et modifiées aux participants, qui décident si la paire est nouvelle (non-étudiée) ou ancienne (étudiée). Ce paradigme évalue les propriétés de la mémoire dite *associative*, un processus essentiel de la mémoire déclarative (Cox & Criss, 2020). Dans le cadre de ces tâches, il a été largement démontré qu'un participant peut apprendre des associations nouvelles dans un temps très limité, généralement sous trois secondes (p ex. Doshier, 1984; Cox & Criss, 2020).

L'encodage d'une association en trois secondes permet d'observer une tendance intéressante : au rappel, une association entre deux items similaires (dits liés, congruents ou cohérents) est toujours mieux reconnue qu'une association arbitraire (dites non-liée, incohérente voire incongruente). Par exemple, Doshier & Rosedale (1991) montrent que les participants reconnaissent les paires cohérentes sémantiquement avec une précision de 88% contre 73% de précision pour les paires arbitraires. Le temps de réaction moyen est

également plus court pour reconnaître les paires cohérentes (RT = 1276 ms) que les paires arbitraires (RT = 1589 ms).

Ces données indiquent que la mémoire sémantique affecte la qualité de l'encodage des informations épisodiques (associations nouvelles en contexte). Plus spécifiquement, il est suggéré que les associations congruentes sont mieux encodées car elles sont soutenues par un réseau sémantique antérieur, qui permet l'encodage d'une trace mnésique plus complexe et plus facilement accessible en mémoire à long terme (Craik & Tulving, 1975; Staresina *et al.*, 2009). A l'inverse, l'encodage et rappel d'associations incongruentes s'appuierait sur un autre type de processus, plus coûteux (Staresina *et al.*, 2009). Ce mécanisme est communément appelé *effet de congruence*, un concept introduit par Schulman (1974) dans ses travaux sur la mémorisation.

Ces réseaux sémantiques antérieurs ont pour effet de causer un second biais de réponse au moment du rappel : quand le participant rencontre une paire cohérente, il aura tendance à catégoriser cette paire comme connue même s'il la rencontre pour la première fois (*false recognition*, en anglais). Cet effet est particulièrement fort quand le participant doit répondre très rapidement, généralement sous une seconde (Doshier, 1984).

De façon très intéressante pour nous, cet effet facilitateur (et parfois trompeur) semble présent peu importe la nature de la caractéristique commune, qu'elle soit sémantique (par ex. Doshier, 1984), orthographique (Cox *et al.*, 2018) ou une caractéristique perceptuelle générale (Cox & Criss, 2020 : 808). Cet effet a également été démontré pour des paires liées par une similarité phonosymbolique, effet que nous allons décrire dans une seconde section.

1.2. Tâche de reconnaissance de paires appliquée au phonosymbolisme

1.2.1. Tâches existantes

De façon analogue aux études évoquées dans la partie précédente, il existe quelques démonstrations empiriques suggérant que le phonosymbolisme offre un avantage cognitif dans des tâches de mémorisation de paires. Quatre études ont, à notre connaissance, testé cet avantage dans une tâche de mémorisation (Aveyard, 2012; Nielsen & Rendall, 2012, Preziosi & Coane, 2017; Sonier *et al.*, 2020). Une tâche classique de reconnaissance de paire liées ou non par une qualité congruente phonosymbolique a par exemple été proposée par Sonier *et al.* (2020), qui montre bien que les paires liées par phonosymbolisme sont mieux reconnues que celles qui ne le sont pas, indiquant ainsi que les considérations théoriques évoquées supra peuvent concerner des schémas phonosymboliques également.

L'épreuve qui nous intéresse le plus a été menée par M. Aveyard (2012). Dans cette étude, quarante participants ont appris 32 paires constituées d'un logatome et une forme visuelle, présentée à l'écran. La moitié de ces paires étaient congruentes au niveau

phonosymbolique et l'autre moitié étaient incongruentes. Dans une première phase, le participant entend un logatome représentatif du pôle {rond} ou {angulaire} et doit choisir qui d'entre deux formes (une ronde, une angulaire) correspond le mieux au logatome entendu. Une fois que le participant a sélectionné sa forme, l'ordinateur lui fait un retour quant à la "bonne réponse", la forme à mémoriser en paire avec le logatome pour le reste de la séance, peu importe ce que le participant a choisi initialement. Le participant apprend successivement les 32 paires. Dans une seconde partie, le participant entend à nouveau chaque logatome rencontré en phase d'apprentissage et doit restituer si le logatome était à mémoriser en paire avec une forme ronde ou angulaire. Les résultats de cette étude montrent que les paires congruentes sont significativement mieux mémorisées que celles incongruentes.

Compte tenu des enjeux théoriques évoqués précédemment, la tâche de mémorisation de paires congruentes et incongruentes sur le plan phonosymbolique inspirée d'Aveyard (2012), semble particulièrement adaptée à la vérification de nos hypothèses relatives à la sensibilité phonosymbolique chez les personnes voyantes et non-voyantes. Effectivement, cette tâche permet d'obtenir à la fois une mesure explicite, à travers les associations produites par les participants, et une mesure implicite, par l'analyse de la précision de mémorisation selon le degré de congruence des paires. L'hypothèse centrale est qu'un participant sensible aux corrélations phonosymboliques retiendra mieux les paires congruentes, car celles-ci bénéficieraient d'un encodage facilité par l'activation de réseaux sémantiques préexistants. À l'inverse, les paires incongruentes mobiliseraient davantage de ressources cognitives, notamment via des processus d'inhibition. Lors du rappel, un biais mnésique en faveur des paires congruentes est ainsi attendu chez les participants présentant un fort sentiment phonosymbolique.

1.2.2. Hypothèses et prédictions

La section suivante présente l'opérationnalisation des hypothèses de l'étude.

- Hypothèse principale (H1) : La congruence phonosymbolique favorise la mémorisation.
- ❖ (H1.1) : La congruence phonosymbolique favorisera la mémorisation dans les groupes, avec des ratios de précision différents selon la condition haptique. Les participants non-voyants seront plus sensibles à l'effet de congruence en modalité de texture, et le groupe voyant en modalité de forme.
 - ❖ (H1.2) : La congruence phonosymbolique favorisera la mémorisation chez les participants voyants par rapport aux groupes non-voyants, en raison d'un mécanisme d'intégration multimodale atypique chez ces derniers.

- Hypothèse complémentaire (H2) : il y aura mémorisation globale des paires.
- Hypothèse complémentaire (H3) : le sentiment phonosymbolique prédira la précision de mémorisation (un sentiment phonosymbolique fort dans un groupe prédit un effet de congruence fort et inversement).
- Hypothèse complémentaire (H4) : le coefficient d'iconicité prédira le sentiment phonosymbolique (plus le coefficient est grand, plus le sentiment est {rond}).
- Hypothèse complémentaire (H5) : les logatomes {ronds/lisses} seront associés à un meilleur sentiment phonosymbolique et une meilleure mémorisation, tel qu'observé dans la littérature (p ex Fort & Schwartz, 2018).

2. Participants

Dans cette partie, nous décrivons le profil de nos populations cibles et participants. Cette étude a visé un recrutement d'un groupe de personnes voyantes et un groupe de personnes non-voyantes. A l'origine, nous souhaitions recruter un groupe non-voyant homogène au profil de déficience visuelle précoce. Des difficultés dans le recrutement ont nécessité de revisiter ce critère. Notre groupe non-voyant est finalement constitué de personnes non-voyantes congénitales et non-voyantes tardives.

2.1. Critères d'inclusion et d'exclusion

Nos critères d'inclusion et d'exclusion pour tous les participants ont été :

- **Âge** : avoir entre 18 et 55 ans.
- **Profil linguistique** : être locuteur du français ou de l'anglais.
- **Profil cognitif** : toute personne présentant une neurodivergence ou trouble cognitif a été exclue de l'étude (notamment les personnes TSA et dyslexiques).
- **Sensibilité haptique** : avoir une perception haptique normale.

Quant au **degré de vue**, comme évoqué supra, des difficultés dans le recrutement ont fait que l'âge de perte de vue de nos participants non-voyants a été obsolète au recrutement : nous avons recruté des participants ayant perdu la vue en fin d'adolescence ainsi que des personnes non-voyantes congénitales. Néanmoins, nous avons contrôlé le degré de vue de chaque participant ainsi que l'âge auquel le participant a atteint un degré de vue correspondant à la cécité sévère.

Le groupe de participants voyants devaient avoir une vue normale ou corrigée et ne pas avoir eu de périodes de malvoyance sévère au cours de leur vie.

2.2. Recrutement

Le recrutement s'est déroulé en deux phases : nous avons d'abord recruté nos participants non-voyants afin d'appareiller le groupe voyant sur leurs profils (notamment, sur l'âge). Pour cela, nous avons contacté une dizaine de structures toulousaines dont le public principal est la population déficiente visuelle. Nos participants ont finalement été recrutés avec l'aide du laboratoire *Cherchons Pour Voir* et via le groupe Facebook *DV_Conseil*. Nos participants ont également diffusé l'annonce de recrutement à leurs connaissances au fur et à mesure des passations. Une fois le groupe non-voyant établi, nous avons cherché à recruter une dizaine de participants voyants. Ces participants ont été recrutés via un questionnaire en ligne diffusé dans les réseaux étudiants et dans des logements de Toulouse.

2.3. Présentation des groupes

Nous avons recruté un total de vingt participants ($n = 20$). Parmi eux, nous avons recruté un total de neuf participants non-voyants ($n = 9$) et onze participants voyants ($n = 11$).

Le profil de chaque participant est résumé dans les tableaux en Annexe 1 et 2. L'historique de degré de vision de nos participants nous a mené à construire trois groupes :

- groupe **non-voyant précoce** ($n = 4$), n'ayant jamais eu accès à la vue.
- groupe **non-voyant tardif** ($n = 5$), ayant perdu la vue après leurs douze ans.
- groupe **voyant** ($n = 11$), ayant toujours eu accès à la vue.

2.3.1. Âge, genre, catégorie socio-démographique et éducation

Au niveau de l'âge, le groupe voyant a un âge moyen de 36,45 ans ($\sigma = 10.7$, Med = 35, Étendue = 21–52). Le groupe non-voyant tardif a un âge moyen de 36,6 ans ($\sigma = 6.67$, Med = 34, Étendue = 29–45). Le groupe non-voyant précoce a un âge moyen de 39,25 ans ($\sigma = 1.79$, Med = 39.5, Étendue = 35–41). Nous avons effectué une ANOVA unidirectionnelle pour petits effectifs, montrant que la répartition de l'âge des participants en fonction de leur groupe est non-significative ($F(2) = .134$, $p = 0.872$). Il y a donc une répartition homogène de l'âge des participants en fonction des groupes.

Au niveau de la répartition de genre, le groupe voyant comprend 6 femmes pour 5 hommes, le groupe non-voyant tardif 1 femme pour 4 hommes, et le groupe non-voyant

précoce 2 femmes pour 2 hommes. Nous avons effectué un test de Fischer pour petits effectifs, montrant que la répartition du genre des participants en fonction de leur groupe est non-significative ($p = 0.497$). Il y a donc une répartition homogène du genre des participants en fonction des groupes.

La majorité de nos participants ont fait des études supérieures (DEUG ou Licence). Les catégories socio-professionnelles de nos participants sont assez hétérogènes. Le détail de la répartition des groupes en fonction de ces critères est résumé ci-dessous.

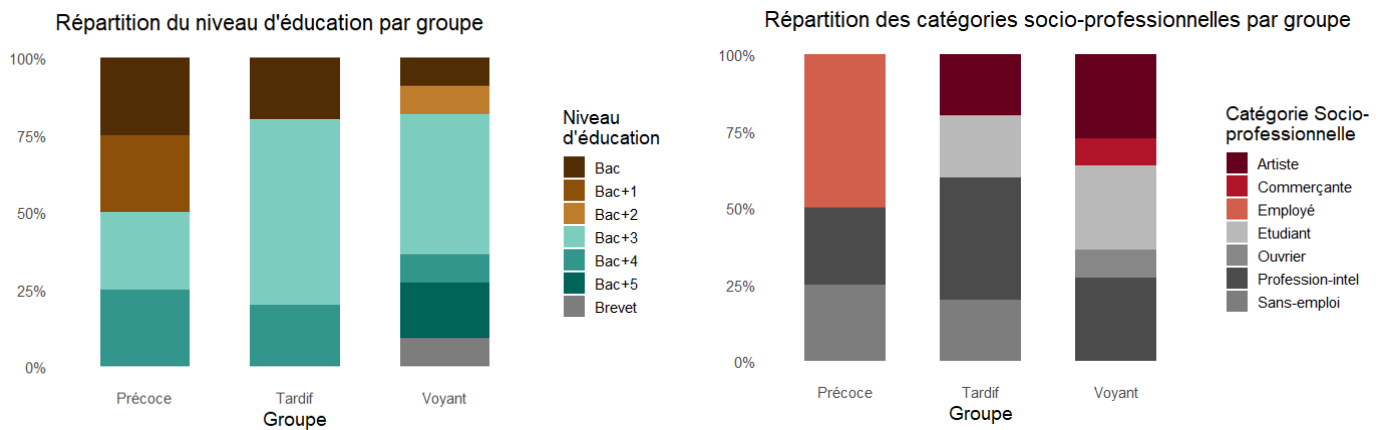


Figure 9. Répartition du niveau d'éducation (gauche) et des catégories socio-professionnelles (droite) selon les groupes.

2.3.2. Âge de perte de vue

Le groupe **voyant** a toujours eu accès à la vue.

Le groupe **non-voyant précoce** n'a jamais eu accès à la vue.

Le groupe **non-voyant tardif** a perdu la vue à un âge moyen de 16 ans ($\sigma = 2.20$, Med = 15, Étendue = 13–19). En moyenne, les participants de ce groupe ont vécu 60.2% de leur vie avec une cécité sévère voire totale ($\sigma = 11.06$, Med = 58%, Étendue = 48%–81%).

3. Matériel

L'objectif de notre étude est de mémoriser deux fois 24 paires constituées d'un logatome et d'un objet haptique. Dans cette partie, nous décrivons donc la création de nos stimuli haptiques et sonores ainsi que les conditions de leur appareillage.

3.1. Items haptiques

Nous avons créé deux groupes de stimuli haptiques : 24 objets haptiques de forme et 24 objets haptiques de texture. Nous justifierons ici l'aspect matériel de ces items.

3.1.1. Formes

3.1.1.1. Considérations d'exploration

On pourrait d'abord se demander si l'exploration haptique d'une forme par une personne déficiente visuelle est analogue à celle faite par une personne voyante. Il a été montré que, dans l'exploration haptique d'une forme, des participants voyants peuvent être significativement moins bons que des personnes déficientes visuelles, mais cet effet est largement dû à des modes d'exploration différents des objets haptiques présentés (Davidson, 1972). Ainsi, il faut éviter ce biais et contraindre les méthodes d'exploration possibles de nos items. Pour cela, nous avons présenté les formes comme des items plats et avons demandé à nos participants de les explorer en les roulant entre le pouce et l'index.

Dans l'étude de Graven & Desebrock (2018) évoquée précédemment, la taille des stimuli de forme était de 30×23 millimètres, ce qui nous semble tout à fait approprié compte tenu que plus un item de forme est petit, mieux il est analysé haptiquement (Gordon & Morison, 1982). Nous avons donc créé des formes dont le contour ne doit pas dépasser une dimension de 3x3 centimètres.

3.1.1.2. Objets de forme

Compte tenu des considérations précédentes, nous avons créé 24 formes (12 rondes, 12 angulaires) en format 2D, matière pâte polymère, qui entrent dans une case de 3x3 cm et mesurent 2 mm d'épaisseur, voir les photographies ci-dessous.

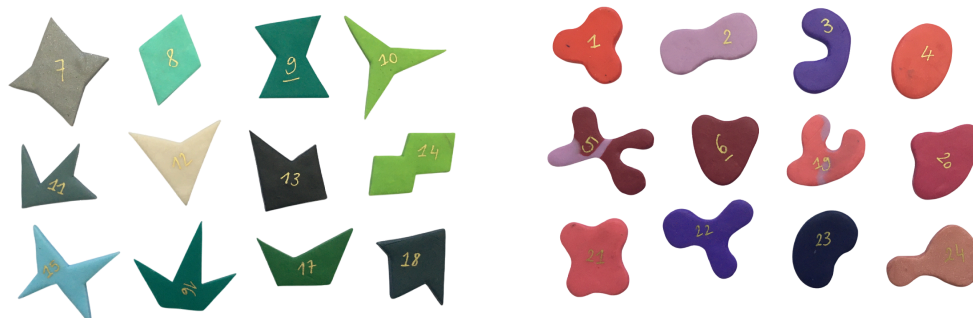


Figure 10. Photographie des 24 items de forme. Le numéro correspond au numéro de la paire. Les items {angulaires} sont à gauche, et les items {ronds} à droite.

Nous avons distribué la géométrie des formes rondes et angulaires de sorte à ce que pour chaque forme ronde, il existe son corrélat angulaire, illustré par les exemple ci-dessous :



Figure 11. Exemple de distribution de la géométrie des objets de formes rondes et angulaires. La forme {ronde} est à gauche et son corrélat {angulaire} à droite.

3.1.2. Textures

3.1.2.1. Considérations d'exploration

On pourrait, de façon analogue aux objets de forme, se demander si l'exploration haptique d'une texture par une personne déficiente visuelle est similaire à celle faite par une personne voyante. Il semblerait que oui : dans leur étude, Baumgartner *et al.* (2015) ont cherché à voir si l'exploration de textures en modalité haptique était perçue d'une façon analogue entre des participants atteint de cécité précoce et des participants voyants travaillant sans voir. Ils ont demandé aux deux groupes d'explorer 70 carrés de matériaux et de caractériser chaque texture sur différentes échelles de qualité de texture, dont le continuum lisse/rugueux. Les deux groupes ont caractérisé chaque texture de façon analogue. A partir de ces observations, on peut conclure que les deux groupes analysent les textures de façon comparable (bien que les mécanismes sous-jacents à cette identification haptique soient peut-être différents), et que cet aspect ne devrait pas poser de problème lors de la passation.

3.1.2.2. Objets de texture

Nous avons créé 24 blocs de textures (12 lisses, 12 rugueux) grâce à des échantillons de différentes textures collées sur des blocs rectangulaires de 9x12 cm de carton mousse, voir les photographies ci-dessous.



Figure 12. Photographie des 24 items de texture. Ordre de numérotation de gauche à droite. Les items {lisses} sont à gauche, et les items {rugueux} à droite.

3.2. Items sonores

3.2.1. Composition phonémique et structure des stimuli

Nous avons créé 48 logatomes distribués en deux groupes de stimuli sonores : 24 logatomes phonosymboliquement représentatifs des pôles {ronds} et {lisses} et 24 logatomes phonosymboliquement {angulaires} et {rugueux}.

Nous avons considéré que les facteurs de congruence phonosymbolique changent selon le groupe (voyant ou déficient visuel). Certains phonèmes n'ont donc pas été utilisés, notamment ceux induisant des représentations phonosymboliques opposées entre les groupes, par exemple la perception du phonème /b/ évoquant la {rondeur} chez le groupe voyant et l' {angularité} chez le groupe non-voyant (Bottini *et al.*, 2018). Cette méthodologie vise à créer des conditions pour favoriser l'égalité des chances à travers les groupes et permettra de lever l'ambiguïté de certains phonèmes en créant deux groupes de logatomes dont la représentation phonosymbolique n'est pas affectée par des différences majeures de perception proprioceptives (articulatoires) ou liées à l'expérience écrite (forme des lettres). A partir de l'étude de Bottini *et al.* (2018), nous avons établi que les phonèmes concernés par cette problématique sont : /v, z, dʒ, b, d, g/ et les avons donc écartés.

Du reste, et compte tenu de la littérature, nous avons considéré que les phonèmes [+continu] et [+voisé] sont perçus comme particulièrement {ronds}. A l'inverse les phonèmes [-continu] et [-voisé] sont perçus comme {angulaires}. Nous avons également considéré que les voyelles [sombres] sont perçues comme plus {rondes} que les voyelles [claires]. Pour former les logatomes, nous avons sélectionné les phonèmes présentés dans le tableau 4, en les considérant comme pluripotentiels (susceptibles d'exprimer à la fois des qualités de forme et de texture).

Qualité phonosymbolique	{rond/lisse}	{angulaire}	{rugueux}
Contraste consonantique	m n l ʃ j w	p t k f s ʃ	p t k f s ʃ ʁ
Contraste vocalique	o u	i e	i e

Tableau 4. Phonèmes sélectionnés pour la création des logatomes de l'étude.

Il a été démontré à plusieurs reprises que les rhotiques expriment particulièrement bien l'impression de rugosité (par ex. Winter *et al.*, 2022). Nous avons ainsi intégré le phonème /ʁ/ dans la configuration des logatomes rugueux.

Sur le modèle de Aveyard (2012), nos logatomes prennent un format trisyllabique (CVCVCV). Afin d'éviter un biais, un logatome ne contient jamais deux fois la même consonne, nous avons donc un modèle consonantique tel que C1.V-C2.V-C3.V. Les consonnes sont réparties de sorte à ce qu'elles apparaissent toutes le même nombre de fois et occupent toutes une position initiale, centrale et finale de façon équivalente. Les voyelles sont

également réparties de sorte à ce qu'elles soient équitablement distribuées entre les items. Un exemple type de logatome {rond} est /monulo/, et /sikepi/ pour un logatome {angulaire}. La liste complète des logatomes est disponible en Annexe 3 et 4.

3.2.2. Validation des logatomes via coefficient d'iconicité

Une fois nos logatomes formés, ils ont été enregistrés avec une voix d'homme, puis une voix de femme (48Hz, 24 bits). Chaque logatome est conservé dans un fichier individuel au format *wav*. d'une durée de 12000 ms, la durée de pause avant le stimuli étant constante. Nous avons ensuite validé ces stimuli sonores grâce au modèle de Fort & Schwartz (2022) en extrayant les caractéristiques acoustiques de chaque item via le logiciel MATLAB (The MathWorks Inc., 2019) à partir du modèle suivant :

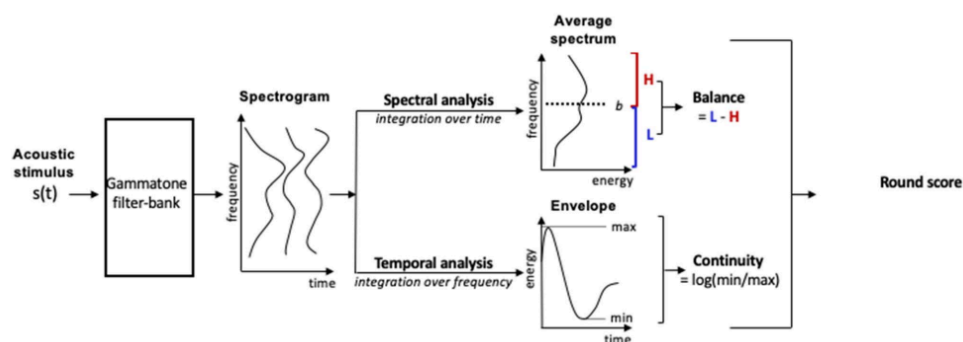


Figure 13. Modèle pour l'obtention du coefficient d'iconicité, issu de Fort & Schwartz (2022 : 2).

Conformément au modèle, nous avons extrait un spectrogramme Gammatone filter-bank pour chaque item. Nous avons ensuite extrait deux mesures :

- (1) La mesure de *continuity* (illustrée par la branche basse de la figure), est une analyse temporelle du signal, calculée comme étant le “ratio entre le maximum et le minimum d'énergie à travers le temps” (2022 : 2). Plus l'énergie est répartie de façon irrégulière, plus le score est bas. Ce score nous permet de classer les items sonores en deux catégories : les logatomes à répartition régulière de l'énergie (qu'on pourrait associer au trait [+continu] ou irrégulière (qu'on peut associer au trait [-continu]).
- (2) La mesure de *balance* (illustrée par la branche haute de la figure) est une analyse spectrale de la répartition de l'énergie dans les différentes bandes de fréquences. Ce score correspond à la différence entre les fréquences basses et hautes. Ici, plus le score est haut, plus les fréquences du signal étudié sont réparties dans les bandes de fréquences graves.

La somme des deux mesures nous donne ce que nous appelons notre coefficient d'iconicité. Dans ce modèle, plus le score est élevé, plus l'item est considéré comme {rond}

ou {lisse}. Plus le score est bas, plus l'item est considéré comme {angulaire} ou {rugueux}, exemplifié par les mesures de quelques stimuli dans le tableau ci-dessous :

Stimulus sonore	Continuité	Continuité moyenne	Balance	Balance moyenne	Coefficient d'iconicité	Coefficient d'iconicité moyen
[lonuwu]	-1,79	-1,615	0,89	0,915	0,9	-0,7
[mulowo]	-1,44		0,94		-0,5	
[piteke]	-4,96	-4,86	0,70	0,635	-4,26	-4,225
[sepeti]	-4,76		0,57		-4,19	

Tableau 5. Scores de continuité, de balance et de coefficient d'iconicité pour quatre logatomes de l'étude, extraites de la voix de femme.

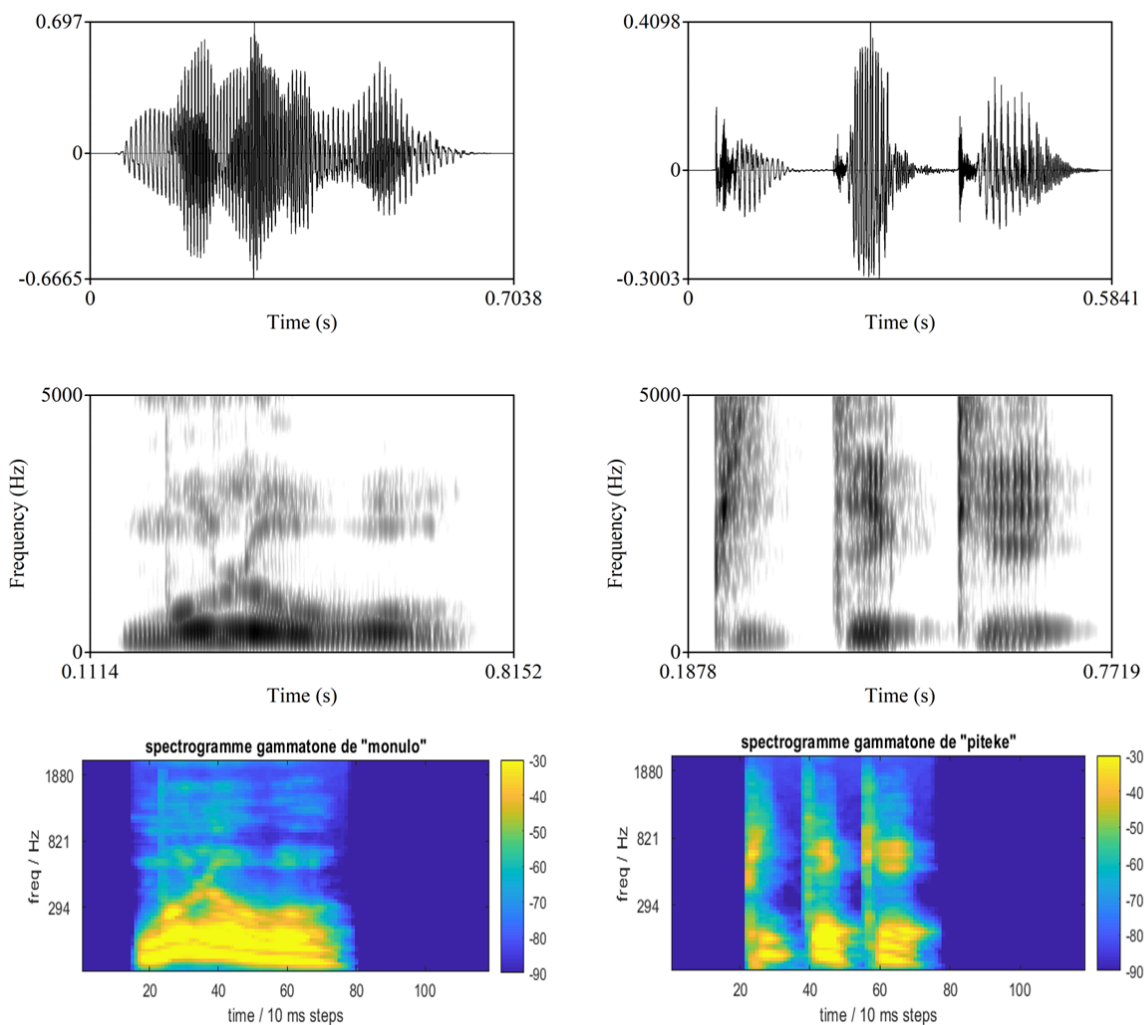


Figure 14. De haut en bas : oscillogrammes, spectrogrammes classiques puis gammatones pour [mulowo] à gauche, [piteke] à droite.

Afin de valider nos items, nous avons cherché à voir dans quelle mesure les scores de *balance* et *continuity* prédisent la qualité phonosymbolique des logatomes (rond/lisse vs

angulaire/rugueux) et déterminer si la performance prédictive et la contribution de ces variables diffèrent selon le genre (Homme vs. Femme).

Une Analyse Discriminante Linéaire (LDA) a été réalisée séparément pour chaque catégorie de la variable Sexe à l'aide de R (packages : *MASS* (Venables & Ripley, 2002), *dplyr* (Wickham *et al.*, 2023), *ggplot2* (Wickham, 2016)). Les scores de Balance et Continuity ont été utilisés comme variables prédictives pour classifier le phonosymbolisme. Les coefficients de la LDA ont classifié les items en fonction du phonosymbolisme avec une précision parfaite (Femme : 100%, Homme : 100%). Ces coefficients ont ensuite été utilisés pour comparer l'importance relative des variables, en tenant compte des différences d'échelle. Pour les items Homme, *balance* était le prédicteur dominant (balance : 4,311213; continuity : 0,257481), tandis que pour les items Femme, les contributions des scores étaient plus équilibrées (balance : 2,095818; continuity : 2,904880).

3.3. Appairage logatomes et items haptiques

Les paires ont été formées de sorte à ce que le nombre de phonèmes et leur emplacement dans l'enchaînement syllabique soit équilibré. Nous avons également équilibré la géométrie des formes et la qualité des textures à travers chaque sous-groupe d'items. A titre d'illustration, voici quelques exemples de paires congruentes et incongruentes sur la base du phonosymbolisme. Toutes les paires sont en Annexe 3 et 4.

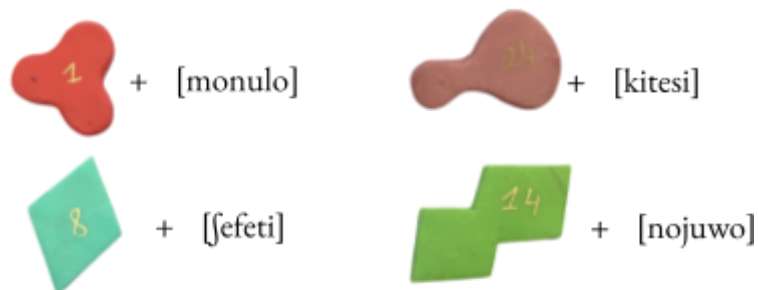


Figure 15. Exemples de paires de forme congruentes (gauche) et incongruentes (droite).

4. Procédure et protocole de passation

4.1. Cadre de passation

Les passations ont eu lieu entre avril et mai 2025. Les séances de passations étaient individuelles, et se sont déroulées en majorité chez le participant (Toulouse et alentours). Sur les vingt passations, cinq se sont déroulées dans les salles des bibliothèques de l'Université et deux chez l'expérimentatrice. Pour créer de l'homogénéité dans leurs expériences physiques, les participants ont tous dû poser les avant-bras sur un tapis.

La séance a généralement duré 1h30. Tous les participants ont participé à la tâche sans voir. Les participants voyants ont passé l'épreuve à travers un mur en carton pour couper leur vue (voir figure 16). Toutes les instructions ont été données à l'oral, les réponses ont été captées avec un micro et les enregistrements ont été perçus à travers un casque. Pour les groupes non-voyants, le formulaire de consentement leur a été lu à l'oral puis transmis par mail au format PDF.



Figure 16. Photographie du cadre de passation en condition de forme, avec le mur de carton pour bloquer la vision aux participants voyants.

4.2. Procédure expérimentale

Dans cette partie, nous décrivons la procédure expérimentale de chaque passation. Le script expérimental utilisé dans l'épreuve a été fait dans *PsychoPy* (Peirce *et al.*, 2019).

Les participants ont eu pour consigne de mémoriser le plus de paires possible et de toujours répondre le plus vite possible. Passé les premiers participants, il leur a été demandé de ne pas essayer de "découvrir" la logique de l'expérience, de se concentrer sur l'apprentissage et que la présentation des paires était aléatoire. Par épreuve, chaque

participant essaie de mémoriser 24 paires constituées d'un logatome plus un objet haptique. La moitié des paires sont congruentes au niveau phonosymbolique et l'autre moitié incongruentes. Un participant passe la même épreuve deux fois : une fois pour tester le phonosymbolisme de forme puis celui de texture. Nous avons équilibré l'ordre de passage à travers les participants (la moitié commencent par la modalité de forme, l'autre en texture).

Mettons, pour la présentation du protocole, que le participant est testé en corrélation de forme. La tâche est divisée en deux phases : une phase d'apprentissage et une phase de mémorisation. Dans une première phase, le participant reçoit deux objets (un dans chaque main) et dispose de cinq secondes pour les explorer avant d'entendre un logatome représentatif du pôle {rond} ou {angulaire}. Le participant doit alors décider, à voix haute, si ce logatome correspond mieux à l'objet rond ou angulaire. A ce stade, on considère qu'il a exprimé son sentiment phonosymbolique.

Une fois que le participant a sélectionné sa forme, il reçoit un retour oral lui indiquant la bonne réponse ("*la paire est à gauche*" ou "*la paire est à droite*"), la forme à mémoriser en paire avec le logatome pour le reste de la séance. Le script tourne indépendamment de la réponse donnée par le participant. Le participant a 3 secondes pour essayer de mémoriser au mieux la paire, avant d'entendre un *clac* qui lui indique de lâcher les objets. L'expérimentatrice lui donne alors deux nouveaux objets et le cycle recommence 24 fois, jusqu'à ce que le participant ait rencontré toutes les paires. Cette phase d'apprentissage dure environ 7'30 minutes et est résumée dans le schéma ci-dessous.

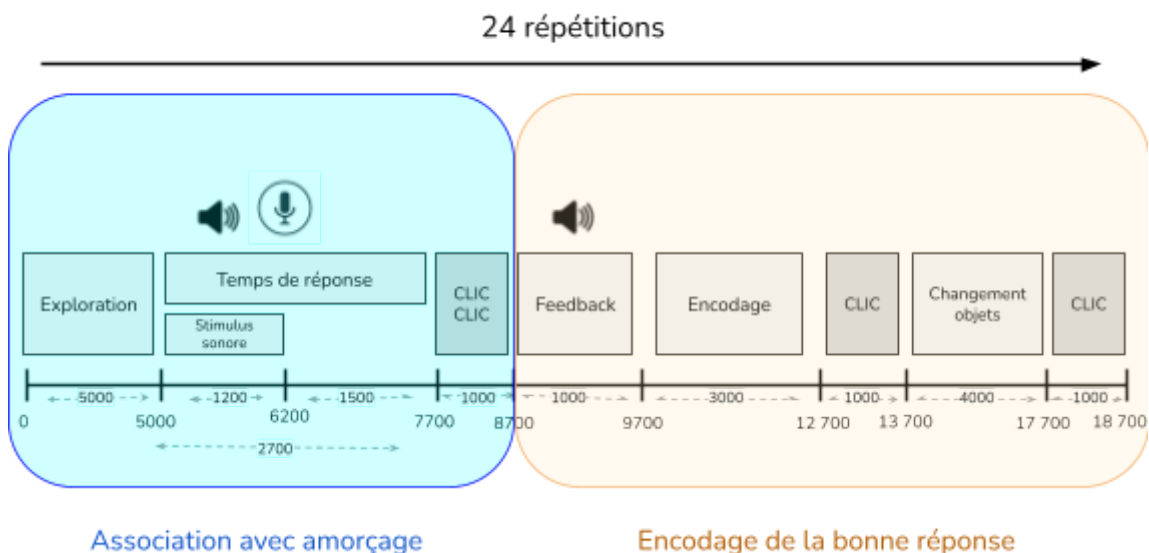


Figure 17. Déroulé de la phase d'apprentissage. Temps en millisecondes.

Une fois la phase d'apprentissage terminée, le participant passe à la phase de rappel. Cette phase est très similaire à la précédente. Le participant reçoit un objet dans chaque main, et l'explore pendant 5 secondes. Il ré-entend un des logatomes précédemment rencontré et

doit reconnaître l'objet auquel le logatome était associé en paire. Pour choisir l'objet, il indique la main gauche ou droite, à l'oral. Le participant restitue ainsi 24 paires. La phase de rappel dure environ 6'05 minutes et est également résumée dans la figure ci-dessous.

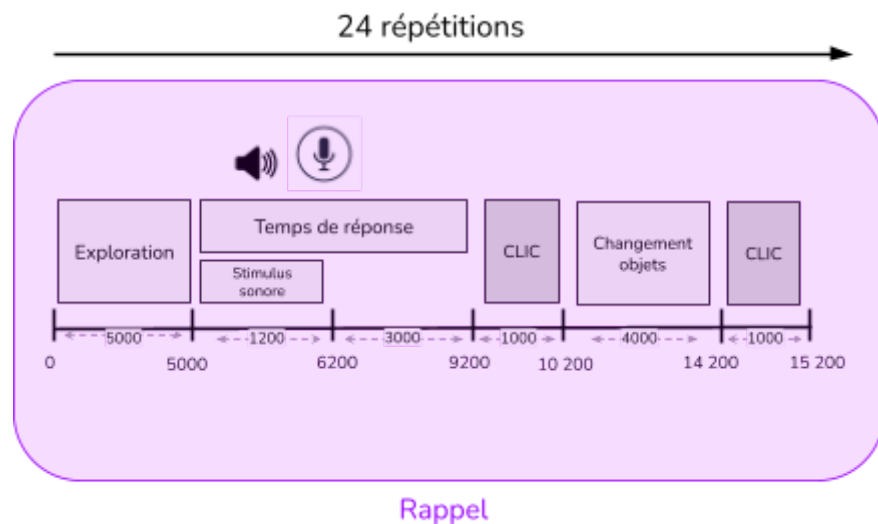


Figure 18. Déroulé de la phase de rappel. Temps en millisecondes.

4.3. Entretien de fin de passation

Une fois les deux épreuves passées, le participant répond à quelques questions :

- Est-ce que certains sons vous ont orienté vers certains objets ? Si oui, lesquels ?
- Etiez-vous plutôt orienté par les voyelles, les consonnes ?
- Quelle était votre stratégie s'il y en a eu une ?
- Avez-vous reconnu certains objets ?
- Une modalité haptique était-elle plus simple que l'autre ?
- Connaissez-vous l'effet bouba-kiki ?

5. Analyse des données

Cette section sera consacrée à l'explicitation des variables, analyses et prétraitements des données de l'étude, en vue de la présentation des résultats. Toutes nos analyses ont été faites dans RStudio (Posit Team, 2025).

5.1. Variables de l'étude

Ci-dessous un tableau récapitulatif des variables de l'étude.

Nom variable	Dép/Indép	Type	Définition
participant	indépendante	nominale	code d'identification du participant
âge	indépendante	continue	âge du participant le jour de la passation
âge_vue	indépendante	continue	âge de perte de vue totale du participant
pourcentage_vue	indépendante	discrète %	pourcentage de vie passé avec une cécité sévère/totale
sexe	indépendante	binaire	genre du participant
CSP	indépendante	nominale	catégorie socio-professionnelle du participant
sous_groupe	indépendante	nominale	groupe auquel appartient le participant
condition_haptique	indépendante	binaire	type d'objet formant les paires (forme/texture)
condition_congruence	indépendante	binaire	association congruente ou non du logatome avec son objet
low_high	indépendante	binaire	groupe phonosymbolique du logatome : plutôt rond/lisse (low) ou angulaire/rugueux (high)
coeff_phono_F	indépendante	continue	somme des score balance et continuité voix femme
coeff_phono_H	indépendante	continue	somme des score balance et continuité voix homme
balance_F	indépendante	continue	score de balance des fichiers voix femme
balance_H	indépendante	continue	score de balance des fichiers voix homme
continuity_F	indépendante	continue	score de continuité des fichiers voix femme
continuity_H	indépendante	continue	score de continuité des fichiers voix homme
phono_accuracy	dépendante	binaire	précision du sentiment phonosymbolique du participant à partir de prédiction de la littérature (0/1)
phono_correct	dépendante	binaire	précision du sentiment phonosymbolique du participant à partir de prédiction de la littérature (correct/incorrect)
mem_score	dépendante	binaire	précision de la mémorisation d'une paire (0/1)
mem_correct	dépendante	binaire	précision de la mémorisation d'une paire (correct/incorrect)
TR_1	dépendante	continue	temps de réaction au sentiment phonosymbolique
TR_2	dépendante	continue	temps de réaction au recall
total_score_mem	dépendante	discrète %	précision de mémorisation d'une paire à travers tous les participants, en pourcentage
total_phono_accuracy	dépendante	discrète %	précision du sentiment phonosymbolique d'une paire à travers tous les participants, en pourcentage
phono_low	dépendante	discrète %	nombre de fois où un logatome à été perçu comme étant {rond/lisse} à travers tous les participants, en pourcentage
mem_low	dépendante	discrète %	nombre de fois où un logatome à été mémorisé comme étant {rond/lisse} à travers tous les participants, en pourcentage

Tableau 6. Variables de l'étude décrites en fonction de leur rôle expérimental, leur type et leur définition de travail.

5.2. Rappel des analyses et interactions testées

Cette section rappelle les analyses et interactions testées dans l'étude.

Pour les hypothèses principales mesurant l'effet de congruence entre les groupes (H1, H1.1 et H1.2), nous avons mesuré :

- la variance de la mémorisation en fonction de la condition de congruence, de la condition haptique et du groupe.
- la variance du temps de réaction au rappel de paires correctement mémorisées, en fonction de la condition de congruence, de la condition haptique et du groupe.

La mémorisation globale (H2) a été mesurée par :

- l'observation de l'évolution de la moyenne de mémorisation au tour d'apprentissage (à priori nul) et au tour de rappel.
- la corrélation entre le choix de l'objet au tour d'apprentissage et de rappel.

La relation entre mesures explicites et implicites de l'étude (H3) a été mesurée par :

- la variance du sentiment phonosymbolique selon la condition haptique et le groupe.
- la variance du temps de réaction au sentiment phonosymbolique, en fonction de la précision de classement phonosymbolique, de la condition haptique et du groupe.
- la corrélation entre le sentiment phonosymbolique et la mémorisation.

La relation entre mesures acoustiques et sentiment phonosymbolique (H4), a été mesurée par :

- la corrélation entre le coefficient d'iconicité et le sentiment phonosymbolique en fonction du groupe et de la condition haptique.
- la corrélation entre la mesure de *balance* et le sentiment phonosymbolique en fonction du groupe et de la condition haptique.
- la corrélation entre la mesure de *continuity* et le sentiment phonosymbolique en fonction du groupe et de la condition haptique.

Enfin, l'effet du type de logatome sur le sentiment phonosymbolique et la mémorisation (H5) a été mesuré par :

- la variance du sentiment phonosymbolique en fonction du type de logatome, de la condition haptique et du groupe.
- la variance de la mémorisation en fonction du type de logatome, de la condition haptique et du groupe.

5.3. Pré-traitement des données

Cette section vise à décrire les prétraitements appliqués au jeu de données de l'étude, afin de justifier l'adéquation de nos modèles et maximiser la validité des résultats.

5.3.1. Modèle binomial pour l'analyse de la variance du score de mémorisation

Dans cette étude, nous avons construit un modèle linéaire mixte généralisé (GLMM). Ce modèle permet d'analyser la variable *score_mem* en fonction de trois facteurs : *condition_haptique*, *sous_groupe*, et *condition_congruence*, ainsi que leurs interactions. Nous avons pris en compte les différences individuelles entre les participants à l'aide d'un effet aléatoire. Ce modèle utilise une distribution binomiale (adaptée aux réponses de type oui/non, succès/échec) et une fonction logit, qui permet de prédire des probabilités.

Avant d'interpréter les résultats, nous avons vérifié que le modèle respecte bien les règles statistiques de base. Pour cela, nous avons utilisé le package *DHARMA* (Hartig, 2024), pour tests de diagnostic. Le test d'uniformité a donné une valeur p de 0.6014, signifiant que les résidus (les différences entre les valeurs observées et celles prédites par le modèle) sont bien répartis. Cela montre que les observations sont indépendantes les unes des autres, une condition importante pour la validité du modèle. Le test de dispersion a donné une valeur p de 0.976, indiquant qu'il n'y a pas de surdispersion. En d'autres termes, la variabilité des données est bien expliquée par le modèle, comme attendu pour une distribution binomiale. Nous avons regardé visuellement les résidus pour chaque participant. Même s'il y a quelques petites variations, rien ne semble indiquer un problème majeur. En résumé, ces vérifications montrent que le modèle est adapté et que nous pouvons faire confiance aux résultats qu'il fournit pour analyser les effets des différents facteurs étudiés.

De plus, l'ANOVA de type III a été utilisée pour obtenir les valeurs de significativité du modèle, notamment car elle permet d'évaluer l'effet de chaque facteur en contrôlant les autres, y compris les interactions et convient bien aux données déséquilibrées comme les nôtres, assurant une interprétation fiable des effets compte tenu de la variabilité entre participants.

5.3.2. Modèle binomial pour l'analyse de la variance du sentiment phonosymbolique

Sur le modèle de la section précédente, nous avons également construit un modèle linéaire mixte généralisé (GLMM) permettant d'analyser la variable *phono_accuracy* en fonction de deux facteurs : *condition_haptique* et *sous_groupe*, ainsi que leurs interactions.

Ici, le test d'uniformité a donné une valeur p de 0.6926, signifiant que les résidus sont bien répartis. Le test de dispersion a donné une valeur p de 0.952, indiquant qu'il n'y a pas de surdispersion. Nous avons également regardé visuellement les résidus pour chaque participant. Même s'il y a quelques petites variations, rien ne semble indiquer un problème majeur. Ces vérifications montrent que le modèle est adapté et que nous pouvons faire confiance aux résultats qu'il fournit pour analyser les effets des différents facteurs étudiés.

Nous avons utilisé, pour les mêmes raisons qu'évoquées précédemment, l'ANOVA de type III pour obtenir les valeurs de significativité du modèle.

5.3.3. Modèle binomial pour l'analyse de la variance en fonction du type de logatome

A l'image des deux modèles précédents, deux modèles linéaires mixtes généralisés (GLMM) ont été construits pour analyser la variance du sentiment phonosymbolique (*phono_accuracy*) et de la mémorisation (*score_mem*) en fonction de trois facteurs : *low_high*, *condition_haptique* et *sous_groupe*, ainsi que leurs interactions.

Le modèle mesurant le sentiment phonosymbolique respecte bien les règles statistiques de base (test d'uniformité : $p = 0.7533$; test de dispersion : $p = 0.848$). C'est également le cas pour le modèle mesurant la mémorisation (test d'uniformité : $p = 0.9374$; test de dispersion : $p = 0.96$). Visuellement, les résidus pour chaque participant ne semblent pas indiquer un problème majeur.

5.3.4. Modèle linéaire pour l'analyse de la variance du temps de réaction

Pour poursuivre, nous avons construit deux modèles linéaires mixtes (LMM) permettant d'analyser les variables de temps de réaction (*TR_1* et *TR_2*) en fonction de trois facteurs respectivement. Nous avons pris en compte les différences individuelles entre les participants à l'aide d'un effet aléatoire. Ce modèle utilise une distribution continue et une fonction logit, qui permet de prédire des probabilités. Nous avons utilisé, pour les mêmes raisons qu'évoquées précédemment, l'ANOVA de type III pour obtenir les valeurs de significativité du modèle.

Les données aberrantes ayant un écart-réduit supérieur à 3 ou à plus de 1,5 écart interquartile de la médiane ont été systématiquement retirées du jeu de données des deux variables (*TR_1* et *TR_2*). Les 960 lignes initiales ont ainsi été réduites à 920. Les lignes contenant des valeurs manquantes (NA) ont également été retirées à l'aide de *complete.cases()*, nous obtenons ainsi 919 lignes au total. Nous avons ensuite vérifié que le modèle respecte bien les règles statistiques de base.

Pour la variable *TR_1*, les facteurs sont : *condition_haptique*, *sous_groupe*, et *phono_correct*, ainsi que leurs interactions. La préparation des données pour l'ajustement du modèle linéaire mixte (LMM) a impliqué plusieurs étapes :

- La normalité des résidus étant significative ($p < .05$), la variable *TR_1* a été transformée en logarithme pour stabiliser la variance et corriger la non-normalité des résidus, ce qui constitue une hypothèse importante pour les modèles LMM. Cette non-normalité a été confirmée par des graphiques Q-Q initiaux. Le test de Shapiro-Wilk confirme une distribution proche de la normalité des données suite à la transformation ($p = 0.0323$).
- Les résultats ne suggèrent pas la présence d'hétéroscédasticité marquée.
- La normalité des effets aléatoires est respectée ($p = 0.2528$).
- L'hypothèse de l'indépendance des observations est respectée. La structure aléatoire (1 | participant) suffit et il n'y a pas de violation de l'hypothèse d'indépendance.

Les mêmes étapes de préparation ont été appliquées pour la variable *TR_2*, dont les facteurs sont : *condition_haptique*, *sous_groupe*, et *condition_congruence*.

- La normalité des résidus étant significative ($p < .05$), la variable *TR_2* a été transformée en logarithme pour stabiliser la variance et corriger la non-normalité des résidus. Cette non-normalité a été confirmée par des graphiques Q-Q initiaux. Le test de Shapiro-Wilk confirme la normalité des données suite à la transformation ($p = 0.4086$).
- Les résultats ne suggèrent pas la présence d'hétéroscédasticité marquée.
- La normalité des effets aléatoires est respectée ($p = 0.3968$).
- L'hypothèse de l'indépendance des observations est respectée. La structure aléatoire (1 | participant) suffit et il n'y a pas de violation de l'hypothèse d'indépendance.

5.3.5. Corrélations entre variables continues

Enfin, des corrélations complémentaires ont été étudiées. Les variables concernées par ces corrélations sont les variables continues *phono_low*, *mem_low*, *total_phono_accuracy*, *total_mem_score* et les variables acoustiques. Nos quatre premières variables sont discrètes, représentent des pourcentages de réussite et ne comprennent aucune valeur aberrante ou manquantes. Nous faisons également le choix de ne pas retirer de données au sein des variables acoustiques, qui représentent des valeurs générées par un modèle contrôlé par l'étude. La distribution de ces données étant toujours non-paramétrique ($p < .05$), nous avons utilisé le **test de Spearman** (ρ).

Chapitre 3 : Résultats

Dans ce chapitre, nous décrivons les résultats de l'expérience présentée au chapitre précédent. Les hypothèses interrogées seront résumées dans chaque section.

1. Résultats H1 – effet de congruence

Dans cette section, nous souhaitons interroger (H1), prédisant que la condition de congruence favorisera la mémorisation en fonction du groupe et de la condition haptique.

1.1. Effet de congruence selon le groupe et la condition haptique

Les mesures intra-groupe montrent qu'il y a bien un effet de congruence peu importe la condition pour le groupe voyant (Forme : $z = 4.805$, $p < 0.0001$, OR = 3.624, $\sigma = 0.971$; Texture : $z = 4.539$, $p < 0.0001$, OR = 3.313, $\sigma = 0.874$) avec une prédiction de précision de mémorisation des paires congruentes de 0.66 en condition de forme et de 0.64 en condition de texture. Il n'y a pas de différence significative entre la précision de mémorisation des paires congruentes ou incongruentes en condition de forme ou de texture (contraste congruence : $z = 0.32$, $p = 0.7524$, OR = 1.09, $\sigma = 0.29$; contraste incongruence : $z = -0.023$, $p = 0.9818$, OR = 0.994, $\sigma = 0.265$). Le groupe voyant a donc la même performance de mémorisation peu importe la condition haptique.

Contrairement au groupe voyant, le groupe non-voyant tardif montre un effet de congruence significatif en condition de texture seulement (Forme : $z = 1.300$, $p = 0.1936$, OR = 1.635, $\sigma = 0.618$; Texture : $z = 5.244$, $p < 0.0001$, OR = 9.429, $\sigma = 4.030$) avec une précision de mémorisation des paires congruentes de 0.55 en condition de forme contre 0.76 en condition de texture. Le groupe non-voyant tardif a une précision de mémorisation des paires congruentes de texture significativement supérieure à celle en condition de forme ($z = -2.519$, $p = 0.0118$, OR = 0.359, $\sigma = 0.146$), suggérant que la performance de mémorisation du groupe non-voyant tardif est affectée par la condition haptique.

Enfin, le groupe non-voyant précoce semble exprimer la tendance inverse : ce groupe montre un effet de congruence significatif en condition de forme seulement (Forme : $z = 3.014$, $p = 0.0026$, OR = 3.667, $\sigma = 1.580$; Texture : $z = -0.414$, $p = 0.6789$, OR = 0.842, $\sigma = 0.349$) avec une précision de mémorisation des paires congruentes de 0.69 en condition de forme contre seulement 0.56 en condition de texture. En condition de texture, le groupe non-voyant précoce a une précision de mémorisation des paires incongruentes significativement supérieure à celle en condition de forme (contraste incongruence : $z =$

-2.226, $p = 0.0260$, OR = 393, $\sigma = 0.165$), suggérant également que la performance de mémorisation du groupe non-voyant précoce est affectée par la condition haptique.

Les mesures inter-groupes de mémorisation montrent qu'il n'y a pas de différence significative entre la précision du groupe voyant et non-voyant précoce en condition de forme (contraste congruence : $z = 0.295$, $p = 0.9533$, OR = 1.113, $\sigma = 0.404$; contraste incongruence : $z = 0.269$, $p = 0.9609$, OR = 1.100, SE = 0.390) suggérant que les deux groupes ont une performance similaire. En condition de texture, il n'y a pas de différence significative entre la précision du groupe voyant et non-voyant tardif (contraste congruence : $z = 1.587$, $p = 0.2512$, OR = 1.768, $\sigma = 0.635$; contraste incongruence : $z = -1.353$, $p = 0.3659$, OR = 0.621, $\sigma = 0.219$) suggérant que les deux groupes ont une performance similaire.

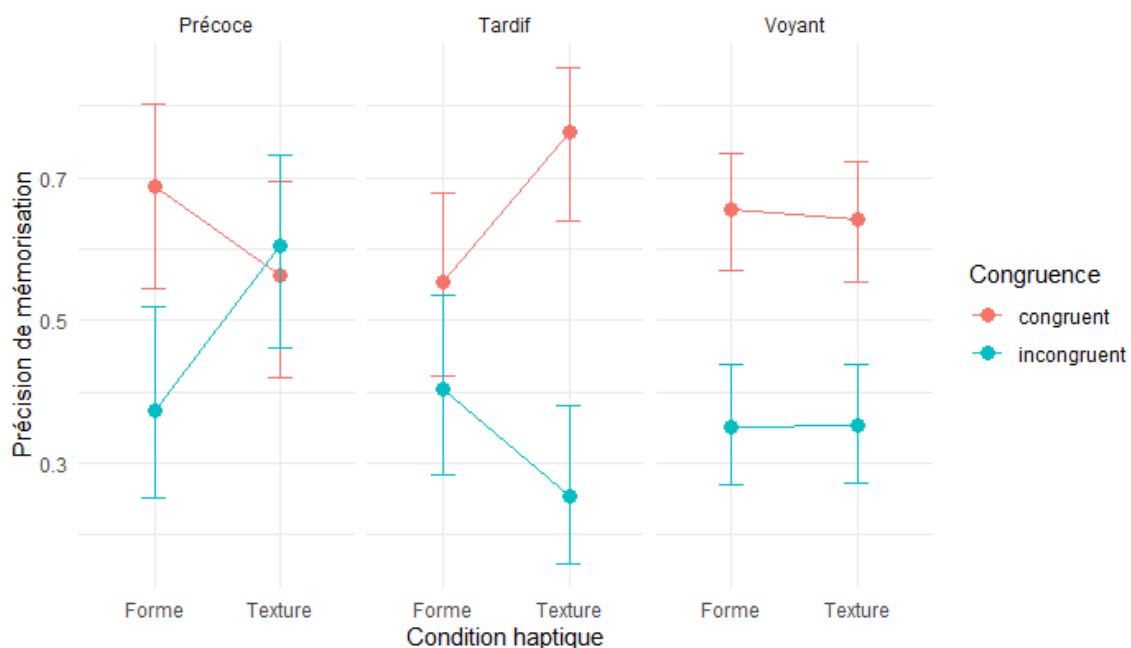


Figure 19. Précision de mémorisation estimée en fonction de la congruence

1.2. Temps de réaction au rappel selon le groupe et la condition haptique

Dans cette section, nous mesurons si les effets de congruence observés supra se reflètent dans les temps de réaction de chaque groupe. Nous nous demanderons si, parmi les paires correctement mémorisées, les réponses aux paires congruentes ont un temps de réaction significativement plus bas que celles des paires incongruentes.

Une visualisation des données (Figure 20) suggère que le temps de réaction pour correctement reconnaître des paires congruentes semble toujours plus court que celui pour correctement reconnaître celles incongruentes, peu importe le groupe ou la condition

haptique. Néanmoins, le seul contraste significatif concerne le groupe voyant en condition de forme ($t = -1.987$, $p = 0.0475$, Estimate = -0.1055 , $\sigma = 0.0531$, $df = 447$).

Les moyennes de temps de réaction estimées selon le groupe⁸ sont les suivantes : chez le groupe voyant (Forme : 2280 vs 2548; Texture : 2262 vs 2364), chez le groupe non-voyant tardif (Forme : 2086 vs 2280; Texture : 1691 vs 1961) et chez le groupe non-voyant précoce (Forme : 1467 vs 1543; Texture : 1657 vs 1773).

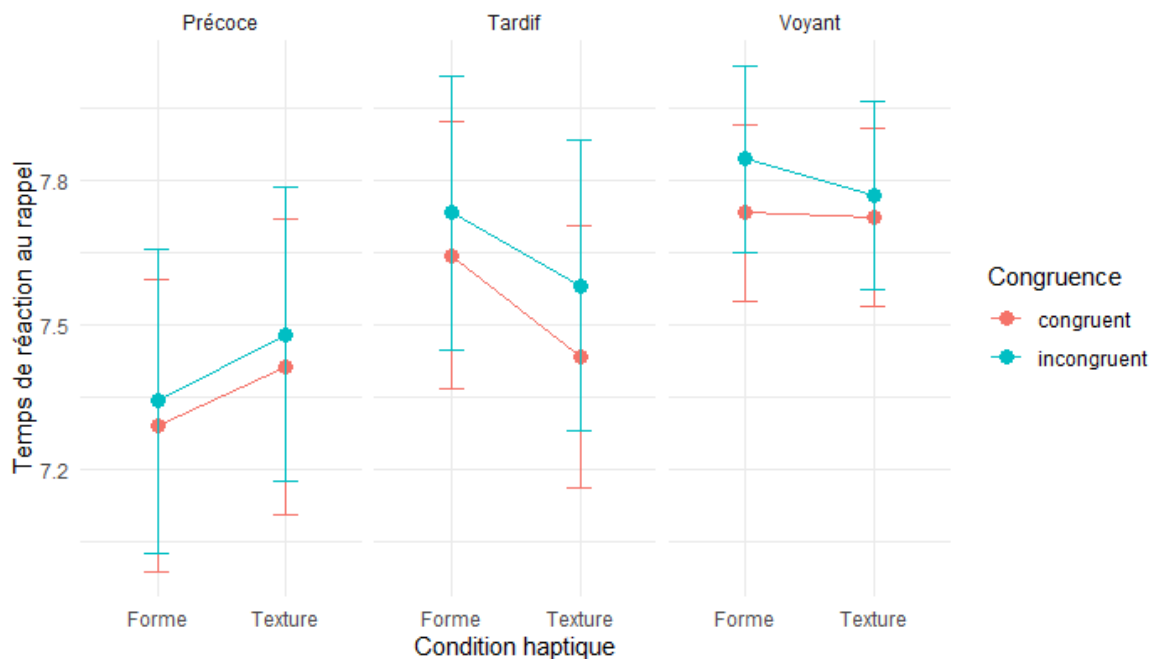


Figure 20. Temps de réaction moyen estimé au rappel de paires correctement mémorisées, en fonction du groupe, de la condition haptique et de la condition de congruence de la paire.

2. Résultats H2 – effet d'apprentissage

Dans cette section, nous souhaitons interroger (H2), supposant qu'il y aura mémorisation des paires lors de la tâche.

D'abord, la mémorisation générale réelle est basse (Moy. tout groupe confondu : 0.51 ; Moy. groupe voyant : 0.51 ; Moy. groupe non-voyant tardif : 0.5 ; Moy. groupe non-voyant précoce : 0.56). Le seul groupe dont la précision de mémorisation augmente entre le tour d'apprentissage (effet de mémorisation à priori nul) et le tour de rappel est le groupe non-voyant précoce (Moy : 0.53 vs 0.56). Pour le groupe voyant (Moy : 0.55 vs 0.51) et le groupe non-voyant tardif (Moy : 0.53 vs 0.5), la précision de mémorisation diminue entre le tour d'apprentissage et le tour de rappel.

⁸ Les temps sont en millisecondes et le temps de réaction en condition congruente est placé en premier.

En lien avec ces premières observations, nous souhaitons voir si le choix de l'objet au premier tour (peu importe la précision du sentiment phonosymbolique) prédit la réponse du participant au tour de rappel. Pour cela, nous mesurons la corrélation entre la perception moyenne d'une paire comme étant {ronde} ou {lisse} et le rappel moyen d'une paire comme étant {ronde} ou {lisse}. Nos résultats montrent que cette corrélation est significative pour le groupe voyant (forme : $n = 24$, $\rho = 0.7603$, $p = 0.00001$; texture : $n = 24$, $\rho = 0.6267$, $p = 0.001$) et le groupe non-voyant tardif (forme : $n = 24$, $\rho = 0.4749$, $p = 0.019$; texture : $n = 24$, $\rho = 0.7685$, $p = 0.00001$). L'effet est fort peu importe la condition pour le groupe voyant ($\rho > 0.5$) et fort en condition de texture pour le groupe non-voyant tardif ($\rho > 0.5$). À l'inverse, cet effet est non-significatif pour le groupe non-voyant précoce (forme : $n = 24$, $\rho = 0.089$, $p = 0.6788$; texture : $n = 24$, $\rho = -0.1439$, $p = 0.5021$).

3. Résultats H3 – corrélation mesures explicites et implicites

Dans cette section, nous souhaitons explorer (H3) supposant que le sentiment phonosymbolique devrait prédire le score de mémorisation. Nous souhaitons d'abord mesurer l'orientation et la précision du sentiment phonosymbolique de chaque groupe, afin d'évaluer si celui-ci peut être corrélé à la qualité de mémorisation.

3.1. Sentiment phonosymbolique selon le groupe et la modalité haptique

Les mesures intra-groupe montrent une différence significative entre le sentiment phonosymbolique de forme et de texture chez les voyants ($z = 3.640$, $p = 0.0003$, OR = 2.11, $\sigma = 0.432$). Les voyants ont un meilleur sentiment phonosymbolique pour la condition de forme, avec un ratio estimé de 0.80 de précision en condition de forme pour 0.66 en condition de texture. Nous observons le même effet pour le groupe non-voyant tardif ($z = 3.643$, $p = 0.0003$, OR = 3.66, $\sigma = 1.300$), avec cette fois-ci un ratio de 0.90 de précision en condition de forme pour 0.70 en condition de texture. Quant au groupe non-voyant précoce, il n'y a aucune différence entre leur sentiment phonosymbolique de forme et de texture ($z = 0$, $p = 1.00$, OR = 1, $\sigma = 0.432$), qui représente 0.61 de précision peu importe la condition.

Il n'y a pas de différence significative entre les sentiments phonosymboliques du groupe voyant et non-voyant tardif, peu importe la condition (forme : $z = 1.686$, $p = 0.2104$, OR = 2.175, $\sigma = 1$; texture : $z = 0.557$, $p = 0.8431$, OR = 1.252, $\sigma = 0.5060$). La seule différence significative entre les sentiments phonosymboliques de nos groupes concerne la condition de forme, où le groupe non-voyant tardif a un sentiment phonosymbolique

significativement meilleur que le groupe non-voyant précoce ($z = -3.229$, $p = 0.0036$, OR = 0.176, $\sigma = 0.0948$). La même différence atteint presque le seuil de significativité entre le groupe non-voyant précoce et voyant ($z = -2.216$, $p = 0.0685$, OR = 0.384, $\sigma = 0.1660$).

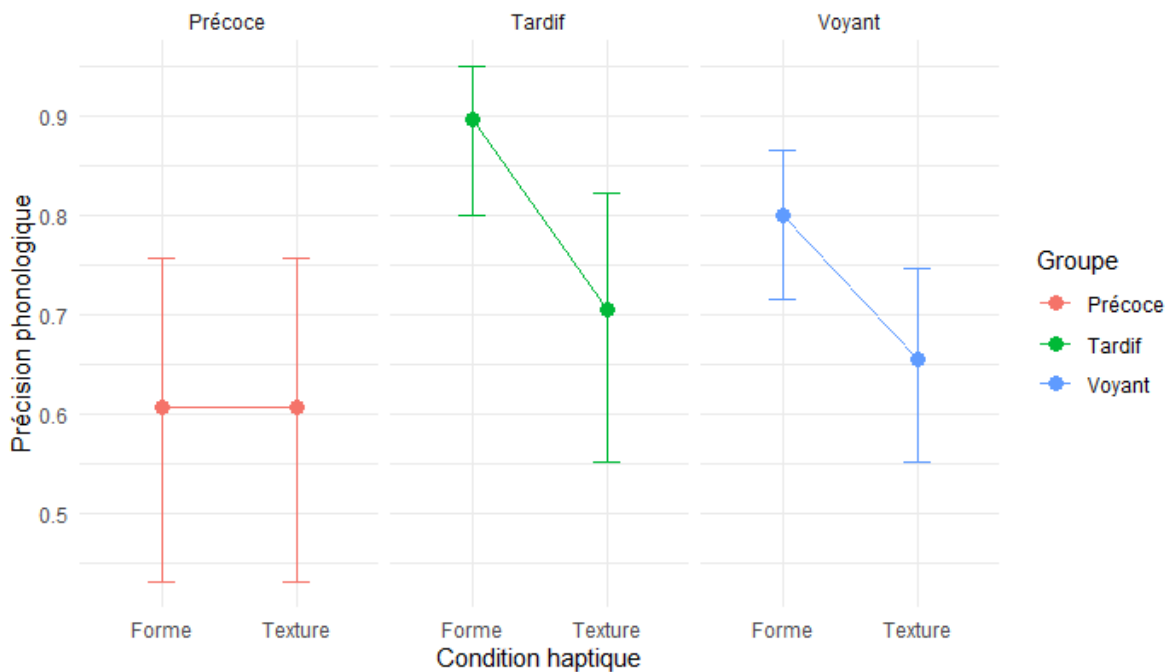


Figure 21. Précision estimée du sentiment phonosymbolique selon le groupe et la condition haptique.

Les sentiments phonosymboliques observés supra se reflètent-ils dans le temps de réaction de chaque groupe ? Pour le groupe voyant, une visualisation des données (Figure 22) suggère que le temps de réaction pour correctement classifier un logatome est toujours plus court que celui pour une classification incorrecte, peu importe la condition haptique. Cette tendance est significative (forme : $t = -4.748$, $p < 0.0001$, Estimate = -0.165, $\sigma = 0.03$, $df = 886$; texture : $t = -2.425$, $p = 0.0155$, Estimate = -0.074, $\sigma = 0.03$, $df = 888$).

Visuellement, cet effet ne se retrouve qu'en condition de texture pour le groupe non-voyant tardif, et en condition de forme pour le groupe non-voyant précoce. Néanmoins, cette différence n'est jamais significative, ce pour le groupe non-voyant tardif ($t = 1.467$, $p = 0.1427$, Estimate = 0.069, $\sigma = 0.05$) ou précoce ($t = -1.112$, $p = 0.2664$, Estimate = -0.051, $\sigma = 0.05$). De façon surprenante, le groupe non-voyant tardif a un temps de réaction plus long pour classifier correctement en condition de texture (correct = 1806 ms; incorrect = 1685 ms).

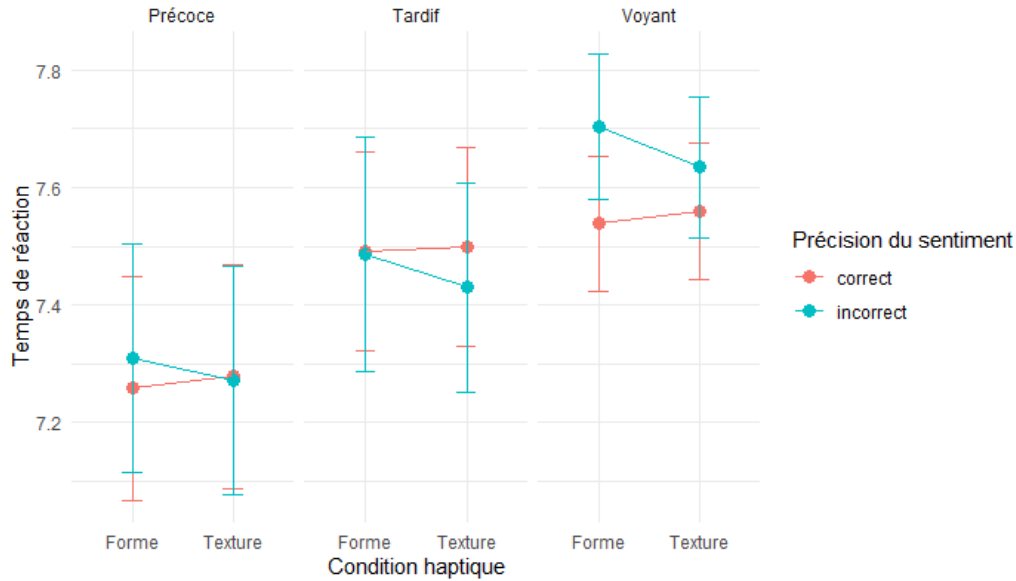


Figure 22. Temps de réaction moyen estimé pour exprimer son sentiment phonosymbolique vis-à-vis d'un logatome en fonction de la précision (correct, incorrect) du sentiment, du groupe et de la condition haptique.

3.2. Corrélation sentiment phonosymbolique et mémorisation

Nous souhaitons maintenant voir si le sentiment phonosymbolique facilite (condition congruente) ou entrave (condition incongruente) la mémorisation. Est-ce que plus le participant perçoit correctement la nature phonosymbolique d'un logatome, plus le biais de congruence est fort ? Pour cela, nous mesurons une corrélation entre la moyenne de précision de sentiment phonosymbolique (*total_phono_accuracy*) et la moyenne de précision de rappel (*total_score_mem*) par paire. Nous évaluons cette corrélation une fois pour les paires congruentes, et une fois pour les paires incongruentes à l'aide d'un filtre.

En filtrage congruent, cette corrélation est significative seulement pour le groupe voyant en condition de forme ($n = 12$, $\rho = 0.6299$, $p = 0.0282$). En filtrage incongruent, cette corrélation est significative seulement pour le groupe voyant en condition de texture ($n = 12$, $\rho = -0.6486$, $p = 0.0225$). Pour le groupe non-voyant tardif et non voyant précoce, l'effet est non-significatif.

4. Résultats H4 et H5 – effet du profil acoustique des logatomes sur la tâche

Dans cette section, nous souhaitons interroger (H4) et (H5), supposant respectivement que le sentiment phonosymbolique peut être prédit par les variables acoustiques issues du

modèle de Fort & Schwartz (2022) et que les logatomes du type {ronds} auront un effet positif sur le sentiment phonosymbolique et la mémorisation.

4.1. Corrélacion sentiment phonosymbolique et coefficient d'iconicité

En premier lieu, nous interrogeons (H4). Les corrélacions entre le coefficient d'iconicité acoustique total et le sentiment phonosymbolique sont significatives pour le groupe voyant (forme_F : $n = 24$, $\rho = 0.7075$, $p = 0.0001$; texture_F : $n = 24$, $\rho = 0.5029$, $p = 0.0122$) et le groupe non-voyant-tardif (forme_F : $n = 24$, $\rho = 0.7206$, $p = 0.00007$; texture_F : $n = 24$, $\rho = 0.7152$, $p = 0.00008$). Le coefficient prédit le sentiment pour ces groupes.

Pour le groupe non-voyant précoce, ces corrélacions ne sont pas significatives ou atteignent presque le seuil de significativité (forme_F : $n = 24$, $\rho = 0.2041$, $p = 0.3388$; texture_F : $n = 24$, $\rho = 0.3948$, $p = 0.0562$). Le sentiment phonosymbolique de ce groupe est mieux prédit par la variable de balance que celle de continuité. Les corrélacions entre le sentiment phonosymbolique et le score de balance sont significatifs ou presque, que ce soit avec la voix de femme ou d'homme (forme_H : $n = 24$, $\rho = 0.3996$, $p = 0.053$; texture_H : $n = 24$, $\rho = 0.4476$, $p = 0.028$). Le score de continuité n'est pas corrélé significativement au sentiment phonosymbolique pour ce groupe, que ce soit avec la voix de femme ou d'homme (forme_F : $n = 24$, $\rho = 0.1712$, $p = 0.4237$; texture_F : $n = 24$, $\rho = 0.3591$, $p = 0.0848$).

Nos résultats montrent qu'il y a une corrélacion significative entre les mesures acoustiques et le sentiment phonosymbolique des participants, peu importe la condition haptique ou le groupe. L'effet semble très fort pour le groupe voyant ($\rho > 0.5$) et le groupe non-voyant tardif ($\rho > 0.5$). Pour le groupe non-voyant précoce, l'effet est plutôt moyen lorsqu'il est significatif ($0.3 > \rho > 0.5$). Bien que tous les participants aient entendu la voix d'homme, les données de la voix de femme prédisent très bien les sentiments.

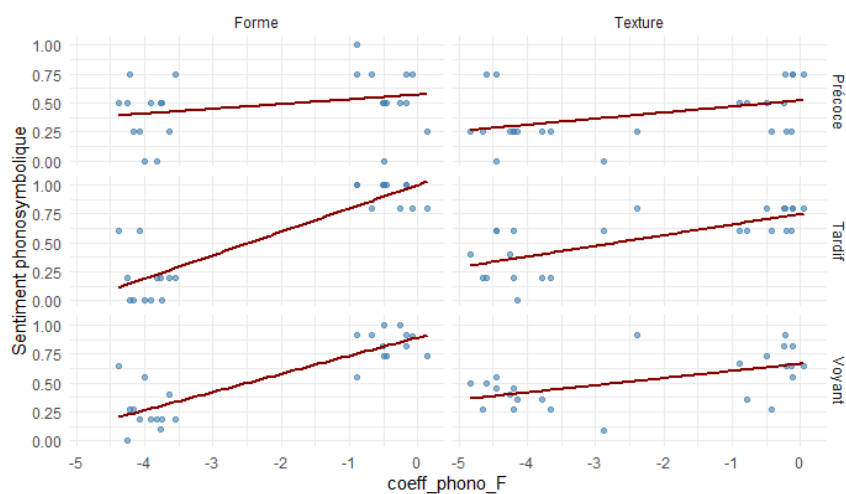


Figure 23. Exemple de corrélacions entre le coefficient phonosymbolique en voix Femme par groupe et condition haptique.

4.2. Effet du type de logatome sur le sentiment phonosymbolique

Dans cette section, nous souhaitons mesurer s'il y a une différence de sentiment phonosymbolique en fonction du type de logatome (*low_high*).

Visuellement (voir figure 24), le groupe voyant a un sentiment phonosymbolique plus précis pour les logatomes {*low*}, mais ce score n'atteint pas la significativité (forme : $z = -1.563$, $p = 0.1180$, OR = 0.604, $\sigma = 0.195$; texture : $z = -0.478$, $p = 0.6329$, OR = 0.875, $\sigma = 0.244$). Le groupe non-voyant tardif a également un sentiment phonosymbolique plus précis pour les logatomes {*low*}, ce score atteint le score de significativité en condition de forme seulement ($z = -2.011$, $p = 0.0443$, OR = 0.242, $\sigma = 0.171$). Le groupe non-voyant précoce a un sentiment phonosymbolique significativement plus précis pour les logatomes {*high*} en condition de texture seulement ($z = 2.472$, $p = 0.0367$, OR = 2.089, $\sigma = 1.070$).

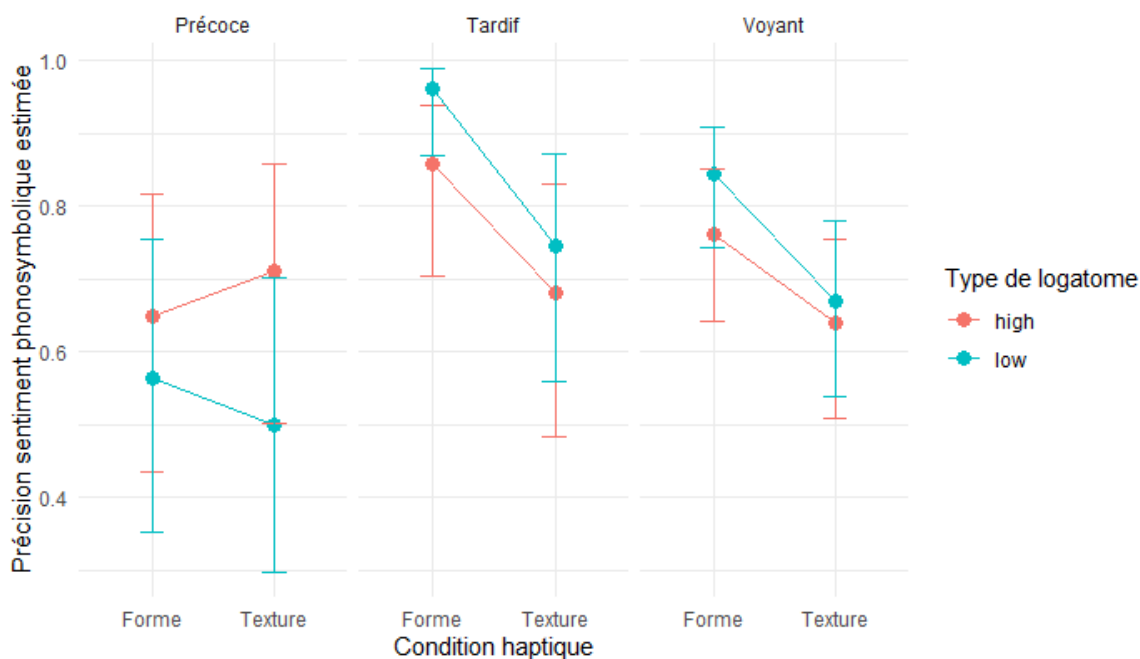


Figure 24. Précision du sentiment phonosymbolique en fonction du type de logatome, du groupe et de la condition haptique.

4.3. Effet du type de logatome sur la mémorisation

Dans cette section, nous souhaitons mesurer s'il y a une différence de mémorisation en fonction du type de logatome (*low_high*).

Visuellement, il semblerait que les groupes non-voyants mémorisent toujours mieux les paires {*high*}. Chez le groupe non-voyant précoce, cette tendance atteint presque la

significativité en condition de forme ($z = 1.830$, $p = 0.0673$, $OR = 2.14$, $\sigma = 0.893$). Pour le groupe non-voyant tardif, cette différence visuelle n’atteint jamais la significativité (forme : $z = 0.730$, $p = 0.4653$, $OR = 1.31$, $\sigma = 0.478$; texture : $z = 0.183$, $p = 0.8551$, $OR = 1.07$, $\sigma = 0.390$). Pour le groupe voyant, il n’y a pas de différence de mémorisation en fonction du type de logatome (forme : $z = 0.061$, $p = 0.9514$, $OR = 1.02$, $\sigma = 0.250$; texture : $z = 0$, $p = 1$, $OR = 1$, $\sigma = 0.246$).

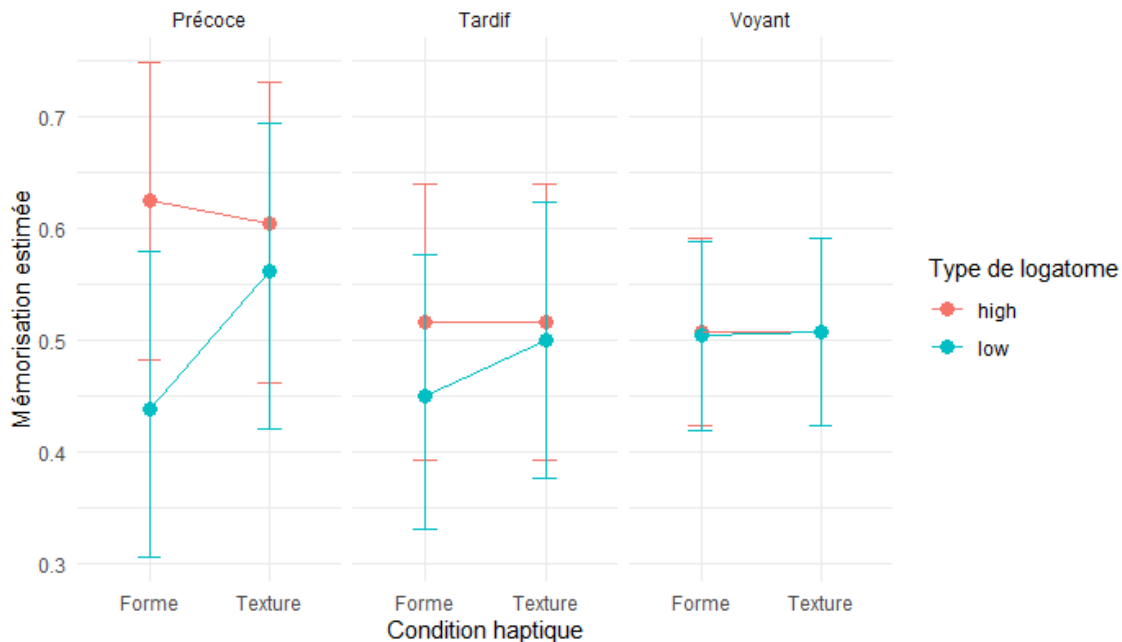


Figure 25. Précision du rappel en fonction du type de logatome, du groupe et de la condition haptique.

5. Retours épilinguistiques sur la tâche

Dans cette section, nous résumons les ressentis de nos participants, recueillis à l’issue de la tâche. La majorité du groupe voyant a trouvé l’exercice en condition de forme plus simple qu’en condition de texture. Cette facilité est généralement non-expliquée, mais quelques participants trouvent que les mots sont “plus appropriés” dans cette condition. 82% des participants ont eu une stratégie liée aux consonnes. Un participant (V14) n’exprime pas d’*effet bouba-kiki*, même lorsque l’expérimentatrice fait passer le test classique avec la vue en fin de tâche. C’est le seul locuteur qui n’a pas grandi dans un environnement francophone (langue maternelle pashto).

Le groupe non-voyant tardif trouve généralement que l'exercice de texture est plus simple, ou ne remarque pas de différence. Deux de nos participants (D03 et D05) ont des tendances synesthètes dans leur quotidien. Ce groupe a eu une stratégie plus concentrée sur les objets rugueux, angulaires et les logatomes associés.

Quant au groupe non-voyant précoce, nous ne notons aucune préférence ou facilité ressentie pour la modalité haptique de forme ou de texture. Nous proposons de restituer le ressenti de chaque participant :

- Le participant D06 rapporte ne pas avoir été orienté vers les objets en fonction des sons. Néanmoins, à partir du second exercice, il explique avoir "*appris*" à associer les mots aux objets. Son instinct phonosymbolique est présent, mais émerge avec le temps d'exposition aux logatomes. Son instinct phonosymbolique est lié aux voyelles (/i/ et /e/, notamment). Ce participant a eu beaucoup d'entraînement au repérage spatial au cours de sa vie.
- La participante D07 rapporte ne pas avoir été orientée vers les objets en fonction des sons. Cet effet était vrai pour l'entièreté de la séance. Cette participante a eu de très grandes difficultés pour se représenter les objets de forme.
- La participante D08 rapporte une tendance phonosymbolique de forme correcte focalisée sur les voyelles. Pour les textures, elle n'a pas exprimé les associations attendues et dit avoir associé "*des mots pas terribles avec des textures pas terribles*".
- Le participant D09 rapporte n'avoir eu aucune perception ou orientation vers certains objets à partir des sons, et nous a par exemple demandé quelle était l'utilité des objets à l'issue du premier exercice. Il rapporte avoir analysé "*le son et l'objet dans des cases complètement différentes*". Sa stratégie a été de se souvenir si le logatome appartenait au concept *gauche* ou *droite* sans se reposer sur les objets.

Chapitre 4 : Discussion et conclusions

1. Interprétation et discussion des résultats H1

Dans cette section, nous interprétons les résultats liés aux effets de congruence dans la tâche, soit (H1). Pour rappel, l'hypothèse centrale est qu'un participant sensible aux corrélations phonosymboliques retiendra mieux les paires congruentes, car celles-ci bénéficieraient d'un encodage facilité par l'activation de réseaux sémantiques antérieurs. Nos principaux résultats sont résumés dans le tableau ci-dessous :

	Corrélations de forme	Corrélations de texture
Voyant	effet de congruence (0.66)	effet de congruence (0.64)
Non-voyant Tardif	pas d'effet (0.55)	effet de congruence (0.76)
Non-voyant Précoce	effet de congruence (0.69)	pas d'effet (0.56)

Tableau 7. Effets de congruence significatifs par groupe et condition haptique, avec précision de mémorisation entre parenthèses.

Nous pouvons d'abord voir que l'effet de congruence est présent chez les voyants peu importe la condition haptique, montrant qu'ils tirent le même avantage cognitif de leurs corrélations phonosymboliques de forme et de texture. Le score de mémorisation étant très similaire dans les deux conditions, on pourrait être tenté de se placer dans la continuité des théories pluripotentiels (p ex. Westbury *et al.*, 2018) suggérant que, chez le sujet typique, différents types de phonosymbolismes sont en fait le reflet d'une seule et même grande association, systématiquement réinvestie dans ce type de tâche. Néanmoins, d'autres mesures comportementales suggèrent que le groupe voyant a un meilleur biais de congruence en condition de forme que de texture, ce qui ne nous permet pas de soutenir pleinement cette possibilité. Notamment, une visualisation des temps de réaction moyen selon le groupe et la condition haptique suggère que le groupe voyant profite particulièrement de l'avantage phonosymbolique en condition de forme (temps de réaction significativement plus court pour reconnaître des paires congruentes, seulement pour ce groupe et en condition de forme).

A l'inverse des voyants, l'effet de congruence émerge selon la condition haptique pour le groupe non-voyant tardif et précoce. Chez le groupe non-voyant tardif, on observe un effet de congruence en condition de texture seulement, démontrant que les personnes non-voyantes tardives tirent un avantage cognitif des corrélations phonosymboliques de

texture, mais pas de celles de forme. Pour le groupe non-voyant précoce, on observe la tendance inverse : un effet de congruence seulement en modalité de forme.

Il est particulièrement intéressant d'évaluer les divergences entre notre groupe contrôle typique (voyant) et les deux autres groupes. Dans un premier temps, l'hypothèse (H1.1), prédisant un effet du changement de paradigmes sensoriels dominants sur les tendances phonosymboliques est partiellement validée par nos résultats, notamment entre le groupe voyant et non-voyant tardif. Effectivement, les contrastes de mémorisation montrent que le groupe voyant a une meilleure précision en condition de forme (0.66 vs 0.55) mais que le groupe non-voyant tardif le surpasse en condition de texture (0.76 vs 0.64), bien que non-significatif. L'effet observé ici découle directement des tendances du groupe non-voyant tardif, qui présente un écart important dans ses scores en fonction de la condition haptique (forme : 0.76; texture : 0.55), alors que le groupe voyant a un score stable à travers les deux conditions (forme : 0.66; texture : 0.64). Ces résultats nous paraissent indicatifs de la réorganisation sensorielle précédemment décrite entre ces groupes.

L'âge avancé de perte de vue du groupe non-voyant tardif (moy. 16 ans) présuppose que les associations multimodales de forme ont été acquises au cours de leur développement, puis recalibrées du fait de leur cécité vers un nouveau domaine pertinent : la matérialité/texture. Ces résultats suggèrent ainsi que le phonosymbolisme est un phénomène évolutif sensible au contexte sensoriel de l'individu percevant. Une dynamique d'acquisition des corrélations phonosymboliques a déjà été décrite dans la littérature du point de vue de la petite enfance (p ex. Fort et *al.*, 2018); notre étude montre néanmoins que cette évolution dépasse les stades développementaux primaires et que les corrélations phonosymboliques peuvent être sujettes à l'évolution jusqu'à l'adolescence.

Cette répartition de la force des corrélations selon le groupe et la condition haptique va dans le sens du modèle de l'intégration multimodale comme fondement au phonosymbolisme. Ici, les deux groupes ont eu une trajectoire développementale marquée par l'accès à la vision dans la petite enfance, et ont donc bénéficié du même point de départ cognitif avant que leurs contextes sensoriels soient modifiés. Dans ce contexte précis, leurs mécanismes d'intégration multimodale seraient donc similaires à l'origine, puis le mécanisme prédictif (aka *coupling prior*; Ernst, 2005) du groupe non-voyant tardif s'est adapté à son nouveau champ perceptif, modifiant par cascade les tendances phonosymboliques au centre de notre étude. Cette théorie est soutenue par le fait que de nouvelles correspondances transmodales prédictives peuvent être apprises en quelques jours à l'âge adulte (Ernst, 2007).

En contraste avec ces premiers résultats, nous remarquons de façon étonnante que les groupes voyants et non-voyants précoces tirent le même avantage cognitif des corrélations phonosymboliques de forme. Contrairement au groupe non-voyant tardif, celui précoce n'exprime pas d'effet de congruence en condition de texture. Cette tendance est

questionnante, notamment car nous avons vu que l'appréhension de la spatialité, nécessaire au traitement des formes, implique des difficultés pour ce public (Hatwell, 2003). De ce fait, nos résultats ne peuvent pas être expliqués par le changement de paradigmes sensoriels dominants entre les groupes.

Étant donné que l'effet documenté chez le groupe non-voyant précoce égale voire dépasse l'effet de congruence observé chez le groupe voyant (0.69 vs 0.66), notre hypothèse (H1.2) prédisant qu'une atypicité dans les processus d'intégration multimodale chez cette population induira par cascade des effets de congruence diminués peu importe la condition est invalidée par nos résultats. Le caractère inattendu de cette tendance complique son interprétation.

A l'oral, les participants non-voyants précoces expriment avoir traité, en majorité, les items de façon unimodale (soit, le logatome n'est pas perçu comme ayant de lien avec l'objet). Cette tendance à traiter les informations multimodales de façon unimodale a été documentée par le passé chez cette population (p ex. Hötting & Röder, 2004). De plus, au moins 2/4 des participants ont abandonné l'exploration des contours des objets de forme lors de la tâche (D07 et D09). Or, l'effet de congruence émerge même lorsque le participant n'arrive pas à développer d'image mentale cohérente des objets de forme (D07). Cela nous amène à penser que ce groupe n'a peut-être pas analysé la forme, mais la magnitude estimée de l'objet (petit/grand) à partir du point de contact de l'objet avec la peau où, contrairement aux formes {rondes} qui sont continues sur la surface cutanée, les piques d'une forme {angulaire} induisent un point de contact sur une petite zone. Dans ce cas, on aurait testé malgré nous un phonosymbolisme de taille, où /píteke/ évoque par exemple la {petitesse} (p ex. Sapir, 1929). Cela serait intéressant, car cette corrélation a l'air active chez ce groupe, qui est sensible aux illusions sensorielles de magnitude (Scheller et al., 2021). Néanmoins, il a explicitement été dit aux participants qu'ils étaient testés sur des objets de forme, et tous ont remarqué l'opposition entre objets {angulaires} et {ronds}. Ces informations invitent à penser que le groupe s'est bien basé sur des informations liées à des qualités de formes.

Si l'effet observé est réellement lié à la modalité phonosymbolique de forme, une autre interprétation est nécessaire. Pour aller plus loin, un participant (D06) explique un développement phonosymbolique similaire à celui observé par Graven & Desebrock (2018) : bien que sa stratégie était unimodale à l'origine, il dit avoir "*appris*" à associer les mots aux objets en cours de tâche, seulement pour la condition de forme. Son apprentissage n'est néanmoins pas arbitraire : il a instinctivement appris les corrélations phonosymboliques typiquement observées dans la littérature (/i/ et /e/ sont {angulaires}). A en croire leur ressenti, le groupe non-voyant précoce développe une sorte de compétence phonosymbolique latente qui n'émerge que pour la condition de forme et qui ne peut pas être expliquée par l'ordre de passage des exercices. Néanmoins, l'effet de congruence en condition de forme est un indicateur clair que cette association n'est pas nouvelle.

La seule différence entre le groupe non-voyant tardif et précoce est l'accès développemental à la vue. La littérature décrit bien qu'il y a atypicité du mécanisme d'intégration multimodale dans les deux groupes du fait de l'absence de vision (ref). Cela étant dit, le groupe non-voyant tardif entame la vie avec un mécanisme typique alors que celui du groupe non-voyant précoce est initialement atypique. Cette différence majeure dans l'implémentation des mécanismes de traitement multimodal chez le groupe non-voyant précoce pourrait induire un mode de traitement phonosymbolique divergent de celui des voyants, quant bien même les résultats suggèrent une performance similaire. De plus, si les deux groupes avaient traité la tâche de façon analogue, le groupe non-voyant précoce aurait été sensible au phonosymbolisme de texture. On pourrait suggérer que la perception des associations phonosymboliques de forme est une prédisposition innée telle qu'il est parfois suggéré dans la littérature (p ex. Ramachandran & Hubbard), et que la vue médie la bonne mise en place de mécanismes de transferts de cette sensibilité à d'autres domaines. Plusieurs études citées dans notre développement entendent que la vue est un sens important quant au fonctionnement typique de la perception multimodale (p ex. Scheller *et al.*, 2021). Dans ce cas, le groupe non-voyant précoce n'aurait pas développé la faculté plastique nécessaire pour recalibrer la capacité phonosymbolique innée de forme vers un domaine haptique qu'il pourrait pourtant mieux aborder.

Pour récapituler, nous avons vu que chaque sous-groupe présente un effet de congruence phonosymbolique dans au moins une modalité, indicateur d'associations phonosymboliques solides peu importe l'accès à la vue. Le groupe voyant a une compétence générale équilibrée à travers les modalités, alors que celles des groupes non-voyants est dépendante de la modalité haptique. Les comportements du groupe non-voyant tardif semblent liés à une réorganisation des paradigmes sensoriels dominants et sont parfaitement expliqués par notre modèle. Néanmoins, les modes de traitement phonosymboliques semblent différents entre le groupe non-voyant tardif et précoce, ce dernier présentant une tendance étonnante, non-prédite par le modèle. A savoir, le groupe non-voyant précoce n'a pas de compétence phonosymbolique diminuée, et égale l'effet de congruence observée chez les voyants en condition de forme. Ici, l'absence de réorganisation phonosymbolique vers le paradigme de traitement dominant pour le groupe non-voyant précoce suggère que ces réorganisations demandent un accès antérieur à la vue.

Toutes les interprétations qui suivent cherchent à trouver d'autres récurrences de ce comportements atypique chez le groupe non-voyant précoce. Y'a-t-il une systématisation des comportements atypiques dans ce groupe ? Si oui, cela irait dans le sens de notre premier avis, supposant que les performances similaires observées entre le groupe voyant et non-voyant précoce sont le résultat de mécanismes divergents.

2. Interprétation et discussion des résultats H2

Dans cette section, nous interprétons les résultats liés à la mémorisation générale au cours de la tâche, soit (H2). D'abord, les résultats montrent que la mémorisation générale réelle est basse et reflète certainement la difficulté de la tâche (Moy. tout groupe confondu : 0.51 ; Moy. groupe voyant : 0.51 ; Moy. groupe non-voyant tardif : 0.5 ; Moy. groupe non-voyant précoce : 0.56).

Pour le groupe voyant (Moy : 0.55 vs 0.51) et le groupe non-voyant tardif (Moy : 0.53 vs 0.5), la précision de mémorisation diminue entre le tour d'apprentissage (effet de mémorisation à priori nul) et le tour de rappel. Ces résultats suggèrent que, à force d'activation répétée des associations phonosymboliques, ces groupes n'arrivent pas à se dégager de l'effet pour mémoriser résultant ainsi en une mémorisation moindre qui diminue pour se rapprocher de 0.5, comme la division des paires congruentes et incongruentes.

A l'inverse, le seul groupe dont la précision de mémorisation augmente entre le tour d'apprentissage et le tour de rappel est le groupe non-voyant précoce (Moy : 0.53 vs 0.56). Par extension avec notre interprétation précédente, ce groupe semble être moins affecté par les associations phonosymboliques que les deux autres dans la mémorisation générale.

Cette divergence entre les groupes voyant/non-voyant tardif et celui non-voyant précoce apparaît aussi ailleurs. Notamment, le groupe non-voyant précoce est le seul pour qui le choix de l'objet au tour d'apprentissage (peu importe la justesse du sentiment phonosymbolique) ne prédit pas le choix de l'objet au tour de rappel, ce qui, pour nous, constitue également une forme de mémorisation.

La corrélation entre la réponse au premier et second tour apparaît pourtant pour le groupe voyant et non-voyant tardif. De plus, on observe ici une tendance similaire à la force de l'effet de congruence : le groupe non-voyant tardif bénéficie d'un effet plus puissant en condition de texture (Non-voyant tardif : $\rho = 0.7685$; Voyant : $\rho = 0.6267$) et le groupe voyant en condition de forme (Voyant : $\rho = 0.7603$; Non-voyant tardif : $\rho = 0.4749$). En termes d'effet général, la force de corrélation en texture pour le groupe non-voyant tardif est égale à celle de forme pour le groupe voyant. Ces résultats suggèrent que les associations phonosymboliques sont plus stables et robustes pour ces groupes que le pour le groupe non-voyant précoce, qui semble pouvoir inhiber l'association plus facilement.

Pour récapituler, il apparaît ici que les trajectoires de mémorisations sont différentes entre les groupes voyants/non-voyant tardifs et le groupe non-voyant précoce, notamment car le groupe non-voyant précoce est : (1) le seul qui s'améliore entre tour apprentissage et tour rappel et (2) le seul pour qui la réponse au premier tour ne prédit pas la réponse au second tour (meilleure inhibition). On remarque encore une fois un traitement particulièrement atypique chez le groupe non-voyant précoce.

3. Interprétation et discussion des résultats H3

Dans cette section, nous interprétons les résultats mesurant un lien entre le sentiment phonosymbolique des participants et leur scores de mémorisation, soit (H3).

	Sentiment de forme	Sentiment de texture
Voyant	0.70	0.66
Non-voyant Tardif	0.90	0.70
Non-voyant Précoce	0.61	0.61

Tableau 8. Précision du sentiment phonosymbolique selon le groupe et la condition haptique.

Pour commencer, tous les groupes expriment explicitement un sentiment phonosymbolique, dans toutes les conditions haptiques. Étonnamment, le groupe ayant le meilleur sentiment phonosymbolique est le groupe non-voyant tardif, qui surpasse légèrement le groupe voyant, bien que non-significatif (forme : 0.90 vs 0.80; texture : 0.70 vs 0.66). Ces résultats sont contraires aux études remarquant une absence d'effet phonosymbolique ou une précision faible chez les personnes non-voyantes tardives, par exemple de 0.55 chez Fryer *et al.* (2014). Peu importe la condition haptique, ce groupe exprime des tendances très similaires voire supérieures au groupe voyant, démontrant qu'ils remarquent les oppositions phonosymboliques de façon très systématique, à hauteur du sujet typique. Contrairement à ce que l'on aurait pu attendre à partir de la répartition des effets de congruence, les deux groupes ont une meilleure conscience du phonosymbolisme de {forme} que de {texture}.

Le groupe non-voyant précoce a, quant à lui, un sentiment phonosymbolique de forme significativement inférieur aux deux autres groupes. Ces participants perçoivent le phonosymbolisme de {forme} et de {texture} à exactement même hauteur (forme : 0.61; texture : 0.61), avec une précision un peu meilleure que le hasard. Cela suggère que ce groupe perçoit seulement quelques régularités phonosymboliques. Ces données sont analogues aux discours de ce groupe qui, à l'oral, ne rapporte pas de sentiment phonosymbolique tranché; 2/4 participants rapportent même une absence totale de sentiment. Ce groupe présente un sentiment réduit par rapport aux autres, confirmant notamment les résultats de Bottini *et al.* (2018) et partiellement l'étude de Sourav *et al.* (2019), montrant par exemple que l'absence de vue affecte négativement l'émergence de mécanismes phonosymboliques conscients chez les personnes non-voyantes seulement si la vue est perdue avant 12 ans. Cette tendance est également similaire à celles observées chez des populations ayant une perception multimodale atypique, à savoir la population sur le spectre autistique ou dyslexique (Oberman & Ramachandran, 2008; Occelli *et al.*, 2013; Drijvers *et al.*, 2015).

Nous avons prédit que la force des sentiments phonosymboliques serait prédictive de la répartition des effets de congruence. Cette hypothèse est invalidée pour les deux groupes non-voyants. A vrai dire, le seul groupe pour qui il y a continuité entre leur sentiment phonosymbolique et l'avantage cognitif qu'ils peuvent en tirer est le groupe voyant, ce dans les deux conditions haptiques. Le biais est positif en condition de forme (plus le sentiment phonosymbolique du logatome est fort, mieux ce logatome est mémorisé en paire avec un objet congruent) et négatif en condition de texture. De plus, le groupe voyant est le seul à avoir un temps de réponse significativement plus court pour correctement classer un logatome au niveau phonosymbolique. Cet effet suggère que le groupe voyant a un sentiment phonosymbolique particulièrement en accord avec leurs associations implicites.

Quant aux groupes non-voyants, leur accès conscient aux corrélations phonosymboliques est incohérent avec la force des associations sous-tendant les effets de congruence. Par exemple, le groupe non-voyant tardif a un sentiment phonosymbolique très précis en condition de forme, mais cette perception ne lui donne aucun avantage dans la tâche de mémorisation. Pour le groupe non-voyant précoce, le sentiment phonosymbolique pour les deux conditions haptiques est similaire, mais seule la condition de forme lui donne un avantage. Cette tendance, démontrée pour la première fois ici, montre que nos participants atypiques n'ont pas accès à leurs propres représentations phonosymboliques et pose question quant à la nature des mécanismes sous-jacents observés.

Ces résultats nous apprennent que, chez un groupe atypique, une conscience phonosymbolique diminuée n'est pas directement liée à une représentation sous-jacente diminuée et inversement. Par extension, cette découverte suggère que la faiblesse de l'*effet boubakiki* observé chez les personnes sur le spectre autistique ou dyslexique n'est pas le reflet d'une compétence généralement diminuée, mais d'un manque d'accès aux représentations phonosymboliques. De façon générale, cette découverte laisse à nouveau entrevoir la limite des tâches d'association en choix contraint (Nielsen & Rendall, 2012).

Pour récapituler, nous avons pu voir que tous les groupes expriment des sentiments phonosymboliques attendus à partir de la littérature. Le groupe ayant la plus grande instance de sentiment est le groupe non-voyant tardif, et celui ayant le sentiment le plus faible est le groupe non-voyant précoce. De plus, l'atypicité sensori-cognitive des groupes non-voyants provoque une incohérence entre leur sentiment phonosymbolique conscient et leurs schémas associatifs internes. A l'inverse, il y a continuité entre les deux pour le public typique.

Bien qu'aucun des deux groupes n'ait accès à leurs représentations internes, le groupe non-voyant tardif se comporte tout de même comme le groupe voyant. Sur ce point, le groupe non-voyant précoce se comporte une fois de plus de façon particulièrement atypique.

3. Interprétation et discussion des résultats H4 & 5

Dans cette section, nous interprétons les résultats mesurant l'effet du profil acoustiques de logatomes sur la tâche. D'abord, nous avons supposé que le coefficient d'iconicité proposé par Fort & Schwartz (2022) prédirait le sentiment phonosymbolique pour tous les groupes, soit (H4). Cette hypothèse est validée chez les groupes voyants et non-voyants tardifs : pour eux, plus un logatome est [aigu] et [discontinu] plus il est {angulaire} ou {rugueux}. Inversement, plus il est [grave] et [continu], plus il est perçu comme étant {rond} ou {lisse}. Ainsi, le groupe non-voyant tardif se comporte, encore une fois, comme un sujet typique.

Contrairement à ces premiers résultats, le coefficient total ne prédit pas le sentiment phonosymbolique du groupe non-voyant précoce. Nonobstant, leur sentiment est prédit par le score de *balance* seul. Des deux mesures constituant le coefficient total, le score de *continuity* ne semble pas pertinent dans la perception phonosymbolique du groupe non-voyant précoce.

Ensuite, nous avons prédit que les logatomes {ronds/lisses} seraient associés à un meilleur sentiment phonosymbolique et une meilleure mémorisation, soit (H5). Cette prédiction est partiellement validée par nos résultats et est, encore une fois, dépendante du groupe. Lors de l'expression de leur sentiment phonosymbolique, les participants voyants et non-voyants tardifs semblent avoir une préférence pour les logatomes {ronds/lisses}. Cette tendance se recoupe avec d'autres tendances observées chez les voyants dans la littérature (p ex. Fort *et al.*, 2018).

Néanmoins, les participants non-voyants précoces présentent la tendance inverse et ont un meilleur sentiment pour les logatomes {angulaire/rugueux}, effet significatif. C'est également le cas au tour de rappel, où cette fois-ci tous les groupes non-voyants semblent mieux mémoriser les paires comprenant des logatomes {angulaires/rugueux}. Cette tendance s'approche de la significativité pour le groupe non-voyant précoce seulement, ce qui est cohérent avec les tendances exposées par ce groupe au niveau du sentiment phonosymbolique. Pour le groupe voyant, la catégorie acoustique du logatome n'affecte pas la mémorisation.

L'effet présenté par les groupes non-voyants est atypique, car il est fréquemment suggéré que les effets phonosymboliques de forme reposent sur un contraste perceptif de type rond/non-rond (Westbury *et al.*, 2018). Pour interpréter ces résultats, on peut envisager que les personnes non-voyantes préfèrent les percepts [discontinus] justement car ils renseignent immédiatement sur le type d'objet rencontré. Lederman & Klatzky (1997) montrent que la discontinuité d'une surface est la première information à être encodée lors de l'analyse haptique vouée à la reconnaissance d'objets. Divers degrés de rugosité ou schémas tactiles saillants seraient donc des fondamentaux dans la représentation haptique générale. A l'inverse, les propriétés spatiales d'un objet (sa forme) sont généralement les derniers

éléments à être perçus. Si le phonosymbolisme découle d'un mécanisme perceptif général voué à l'appréhension optimale de son environnement, la préférence des groupes non-voyants envers les paires {angulaires/rugueuses} semble répondre à leur niche écologique. Les voyants n'ayant généralement pas utilité de reconnaître un objet seulement au toucher, cette tendance n'émerge pas chez eux.

Ces résultats sont étonnants car, on le rappelle, le score de *continuity* d'un logatome ne prédit pas le sentiment phonosymbolique des participants non-voyant précoces, qui sont pourtant les plus sensibles aux logatomes marqués par une discontinuité forte. Le comportement du groupe non-voyant tardif est également étonnant, car ce groupe se comporte ici comme un voyant au premier tour, et comme le groupe non-voyant précoce au second tour. Lorsque ce groupe est affecté par des perceptions liées aux formes, il perçoit mieux les logatomes {ronds}, et mieux les logatomes {rugueux} lorsqu'il est affecté par des perceptions liées à la matérialité. Le comportement de ce groupe montre que la nature de la tâche active différentes associations et modes de perception chez eux.

4. Récapitulatif

A travers les résultats de l'expérience, nous avons pu voir que deux tendances se dégagent. D'abord, le groupe non-voyant tardif a tendance à se comporter comme le groupe voyant, laissant supposer qu'ils ont effectué la tâche via les mêmes mécanismes, dépendant d'un accès développemental à la vue. Ces deux groupes ont par exemple les mêmes tendances phonosymboliques explicites, la même trajectoire de mémorisation, la même sensibilité aux mesures acoustiques et aux logatomes {ronds/lisses} lors de l'expression de leur sentiment phonosymbolique.

A l'inverse, le groupe non-voyant précoce présente souvent un pattern unique, notamment des tendances phonosymboliques explicites faibles et non-influencées par la modalité haptique, une trajectoire de mémorisation marquée par une meilleure inhibition, une sensibilité aux mesures acoustique seulement prédite par la *balance* et une préférence pour les logatomes {angulaires/rugueux} peu importe la nature de la tâche. Leur discours épilinguistique sur la tâche est également différent des deux autres groupes, et les participants expriment être moins sensibles à l'interprétation multimodale.

Malgré cela, les groupes non-voyants présentent quelques points communs. Notamment, ces deux groupes ne semblent pas avoir accès à leurs propres représentations phonosymboliques. De plus, les effets de congruence observés chez eux sont dépendants de la modalité haptique, et ils présentent tous une préférence pour les paires

{angulaires/rugueuses} lors du rappel, interprété comme étant le reflet de l'importance fonctionnelle de la saillance haptique pour ce groupe (Lederman & Klatzky, 1997).

Certains de ces résultats vont dans le sens de la littérature, notamment l'effet phonosymbolique explicite faible chez le groupe non-voyant précoce (p ex. Fryer *et al.*, 2014) ainsi que leur interprétation d'apparence unimodale (p ex. Hötting & Röder, 2004). D'autres résultats contredisent la littérature : ici, le groupe non-voyant tardif présente un sentiment phonosymbolique très fort, et le groupe non-voyant précoce présente une association implicite de forme aussi forte qu'un voyant, effet qui n'avait jamais été documenté par le passé.

La mesure qui nous intéresse le plus ici était la répartition des effets de congruence lors du rappel, selon la modalité haptique. Nos résultats montrent d'abord que les voyants profitent de cet effet dans les deux modalités, mais sont légèrement plus sensibles en condition de forme. Celle-ci étant la correspondance phonosymbolique la mieux documentée, pour laquelle nous avons des preuves empiriques nombreuses de sa présence à un très jeune âge (p ex. Fort *et al.*, 2018) et de façon universelle (p ex. Ćwiek *et al.*, 2021), nous la considérons comme la correspondance de référence, la plus forte chez les sujets typiques. La correspondance de {forme} serait importante pour que les voyants forment de nouvelles associations dont, potentiellement, celles de {texture} (Hamilton-Fletcher *et al.*, 2018). Cette correspondance originelle, similaire en beaucoup de points aux corrélations phonosymboliques de {taille}⁹, permettrait à ce groupe de développer des associations phonosymboliques dans de nombreux autres domaines, comme celui du {goût}, de la {luminosité}, de la {valence} etc, pourvu que le groupe percevant ait eu accès à la vue dans son développement.

Effectivement, l'accès développemental à la vue semble conditionner l'habileté à transférer les associations phonosymboliques vers un nouveau domaine sensoriel pertinent. Le groupe non-voyant tardif exemplifie parfaitement cette théorie en présentant un effet de congruence surpassant celui des voyants en condition de {texture}, répondant ainsi à leur niche écologique tel qu'il avait été prédit par notre modèle. A l'inverse, le groupe non-voyant précoce ne présente pas cette tendance, et n'est sensible qu'au phonosymbolisme de {forme} (ou de {taille}, selon l'effet véritablement mesuré), ce qui ne répond pas à leur niche écologique. Leurs données comportementales vont à l'encontre du modèle, notamment car la sensibilité des personnes non-voyantes précoces aux correspondances transmodales simples de texture n'a pas prédit la sensibilité au phonosymbolisme de texture (Hamilton-Fletcher *et*

⁹ Étant donné qu'une forme {angulaire}, pour le même périmètre, occupe une plus petite surface qu'une forme {ronde}, ces deux correspondances semblent se répondre (Kac, 1966; Fort & Schwartz, 2022). Ohala (1994) proposait déjà que les correspondances de {taille} sont phylogénétiquement ancrées dans notre cognition car ont une implication éthologique. Il observe notamment que les animaux imitent des changements de taille physique à travers des modulations sonores, indiquant la soumission avec des signaux aigus pour manifester la {petitesse}, et des signaux {graves} pour manifester une grande taille, facteur de dominance. Ces phénomènes seraient encore présents chez l'humain dans les attitudes de politesse (voix plus douce) et de colère (voix plus grave) et auraient une caractéristique innée justement du fait de leur importance dans la survie sociale.

al., 2018). Ce groupe est-il sensible aux correspondances de {forme/taille} car elles représentent la correspondance cognitive de base, mais n'ont pas les ressources pour faire des transferts phonosymboliques du fait d'une privation visuelle précoce ? La question du transfert sensoriel chez les personnes non-voyantes précoces est un débat qui a traversé les siècles et renvoie au problème de l'aveugle de Molyneux (Locke, 1689). A l'origine, les philosophes William Molyneux et John Locke se demandaient si, une fois la vue recouvrée, une personne aveugle pourrait reconnaître par la vue des objets connus par le toucher. Held *et al.* (2011) montrent que ce transfert sensoriel n'est pas immédiat chez les patients post-cataracte. Cette habileté s'établit dans le temps, quand bien même les participants suivis sont très jeunes (8-17 ans). Les mécanismes sous-tendant les transferts intermodaux ne sont donc pas innés, mais acquis par expérience. Compte tenu de cela, il est intéressant de se demander si l'absence développementale de vision affecte la mise en place du mécanisme qui permet au groupe non-voyant tardif de développer de très bonnes correspondances transmodales de texture.

5. Limites de l'étude

Les résultats de notre étude se confrontent à plusieurs limites méthodologiques qu'il est important de citer. Tout d'abord, la plus grande limite est que les groupes non-voyants sont constitués de peu d'effectifs. L'effectif réduit diminue la probabilité que nos résultats soient extensibles à la population générale. Nous avons pu pallier ce problème avec nos modèles statistiques, mais augmenter le nombre de participants permettrait d'obtenir des résultats plus représentatifs.

Quant à la préparation des items, nous avons estimé que les phonèmes représentatifs des pôles {rond~angulaire} sont pluripotentiels et peuvent s'adapter aux paradigmes de texture. Bien que cela semble vrai pour les groupes voyants et non-voyants tardifs, peut-être ne le sont-ils pas pour le groupe non-voyant précoce, expliquant ainsi leurs tendances atypiques. A titre d'exploration, nous avons mesuré informellement si les tendances phonosymboliques du groupe non-voyant précoce en condition de texture peuvent être prédites par le nombre moyen de points brailles nécessaires à l'écriture des logatomes, supposant que plus un logatome est dense, plus il serait perçu comme {rugueux}. Cette piste n'a rien donné, mais questionner le sentiment phonosymbolique des personnes non-voyantes sur le phonosymbolisme de {texture} en isolant les phonèmes serait intéressant.

Ensuite, nous n'avons pas établi un protocole de validation systématique et précis des objets haptiques de forme et texture. Faire valider les items par une population non-voyante et voyante avant la tâche aurait été intéressant.

Lors de la passation, les premiers participants n'ont pas reçu la consigne explicite de ne pas chercher la logique de l'épreuve et ont mis en place des stratégies non-contrôlées par l'étude. Enfin, l'encodage des données (1920 enregistrements de quelques secondes) a été fait à la main et laisse une place possible à l'erreur humaine, bien que peu probable.

Du point de vue théorique, notre étude invite à exploiter davantage certains axes qui ont consciemment été écartés ici. Notamment, nous n'avions pas l'ambition d'interroger la littérature neuro-anatomique autour des mécanismes perceptifs abordés ici, qui pourraient pourtant éclaircir des aspects importants de notre développement.

Conclusion

Dans cette étude, nous nous sommes intéressées à l'impact de la cécité sur les mécanismes phonosymboliques typiques. Nous avons vu que la littérature antérieure montrait déjà que les personnes non-voyantes présentent un profil atypique en matière de perception phonosymbolique (Fryer *et al.*, 2014; Bottini *et al.*, 2018), souvent comparable à celui observé chez les personnes atteintes de dyslexie ou de troubles du spectre autistique (Drijvers *et al.*, 2015 ; Oberman & Ramachandran, 2008). Certaines études suggèrent que les correspondances transmodales pourraient émerger de mécanismes d'intégration multimodale (par ex. Ernst, 2007), et qu'à ce titre, elles seraient sensibles aux configurations écologiques propres aux profils de perception atypiques (par ex. Bottini *et al.*, 2018 ; Hamilton-Fletcher *et al.*, 2018). Pour explorer cette hypothèse, nous avons conçu une tâche de mémorisation permettant d'évaluer le rôle potentiel du phonosymbolisme dans la facilitation cognitive de la mémorisation de paires de mots congruentes et incongruentes d'un point de vue phonosymbolique. Deux modalités perceptives ont été testées : la forme (ronde vs angulaire) et la texture (lisse vs rugueuse), auprès de trois groupes de participants : un groupe voyant, non-voyants devenus et un groupe de non-voyants congénitaux.

Après exploration de nos résultats, revenons à nos interrogations initiales : la vue est-elle nécessaire à l'émergence du phonosymbolisme ? Pourquoi les personnes non-voyantes n'expriment-elles pas les corrélations phonosymboliques de façon analogue aux voyants ? Au terme de ce travail, nous pouvons conclure que la vue n'est pas nécessaire à l'émergence de corrélations phonosymboliques profondes; son absence a cependant des

conséquences notables. D'après notre étude, l'accès ininterrompu à la vue permet aux voyants de développer des correspondances phonosymboliques profondes indépendamment du domaine sensoriel testé. En plus de cela, les voyants bénéficient d'une cohérence sensorielle entre les corrélations phonosymboliques perçues au niveau conscient et celles implicites. En parallèle, l'absence de vision ne permet pas de maintenir (voire développer) toutes les correspondances phonosymboliques auxquelles sont sensibles les voyants. Notre étude montre très clairement que la privation visuelle prolongée induit une restructuration des tendances phonosymboliques chez un sujet, dépendant de l'âge de perte de vue.

D'abord, un accès développemental à la vue (non-voyance tardive) semble plutôt provoquer une réorganisation des paradigmes phonosymboliques dominants, répondant à la niche écologique du sujet. Les personnes non-voyantes tardives profitent d'un avantage phonosymbolique lié aux propriétés tactiles matérielles de la {texture}, le domaine de spécialité de l'exploration haptique qui pallie bien, chez ce groupe, l'absence de vision (Hatwell, 2003).

Ensuite, la privation visuelle congénitale semble provoquer une plus grande atypicité dans les tendances phonosymboliques. Même si l'on observe bien une réorganisation des paradigmes phonosymboliques chez les personnes non-voyantes tardives, leur comportement se rapproche de celui des voyants à plusieurs égards. A l'inverse, le seul point commun entre le groupe voyant et non-voyant précoce est une sensibilité prononcée pour les corrélations phonosymboliques de forme. Cela nous invite à penser que la privation visuelle congénitale induit un mode de traitement phonosymbolique différent de la privation tardive. Nous avons suggéré qu'une mise en place atypique de la perception multimodale chez cette population ne permettrait pas de faire les transferts intermodaux qui paraissent sous-tendre la diversité des domaines phonosymboliques observables chez les personnes voyantes ou le transfert vers la dominante {texture} chez le groupe non-voyant tardif.

A l'origine, nous avons émis l'hypothèse que les données de l'étude seraient prédites par les configurations divergentes du mécanisme prédictif d'intégration multimodale selon le groupe. Ici, cette prédiction est avérée pour le groupe ayant eu accès à la vue dans l'enfance (non-voyant tardif). Ce modèle doit être approfondi pour intégrer le comportement inattendu du groupe non-voyant précoce, pour qui l'absence de vue semble altérer la perception phonosymbolique, sans pour autant la diminuer ou la transférer vers leur domaine écologique qu'est, à priori, la matérialité. Plus d'études doivent être faites sur ce sujet avant d'advenir à une conclusion claire.

Cette étude ouvre pour nous plus de questions qu'elle n'en clos. Elle propose toutefois des apports intéressants que nous citerons ici. Généralement, ce projet a permis de nuancer les résultats hétérogènes de la littérature, notamment ceux publiés par Bottini *et al.* (2018), qui supposaient déjà que les tendances phonosymboliques des personnes non-voyantes n'étaient pas seulement diminuées, mais réorganisées. A notre connaissance, ce projet est le premier à mesurer les corrélations phonosymboliques de façon implicite chez une population

atypique. Ce faisant, nous avons pu montrer l'étendue de la divergence comportementale entre les groupes déficients visuels précoces et tardifs au niveau phonosymbolique, interrogeant ainsi l'impact d'un accès développemental à la vue sur la perception multimodale. Ce projet a également permis de démontrer que les effets phonosymboliques sont plastiques jusqu'à au moins la fin de l'adolescence. Enfin, cette étude a permis de découvrir qu'il existe une discordance entre le sentiment phonosymbolique explicite et implicite chez le public non-voyant, permettant de supposer que cet effet s'étend également à d'autres populations atypiques présentant un effet phonosymbolique diminué.

Cette étude mériterait d'être reproduite avec plus de participants pour valider ou invalider nos premiers résultats. Quant à l'interprétation des données, nous aurions certainement besoin d'une meilleure vue d'ensemble sur les comportements multimodaux (intégration, transferts) chez la population non-voyante. Nous n'avons par exemple pas d'explication quant à l'incohérence observée entre le sentiment phonosymbolique perçu et réel des participants non-voyants. Nous estimons que les tendances observées ici ne peuvent pas être expliquées par une étude comportementale seule : des données neuro-électriques pourraient certainement nous offrir des apports édifiants.

Bibliographie

Arias, C. (1996). L'écholocalisation humaine chez les handicapés visuels. *L'année psychologique*, 96(4), 703–721.

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory & Cognition*, 40(1), 83–92.

Baumgartner, E., Wiebel, C. B., & Gegenfurtner, K. R. (2015). A comparison of haptic material perception in blind and sighted individuals. *Vision research*, 115(Pt B), 238–245.

Bottini, R., Barilari, M., & Collignon, O. (2018). Sound symbolism in sighted and blind. The role of vision and orthography in sound-shape correspondences. *Cognition*, 185, 62–70.

Boudier, F. (2018). Les recherches expérimentales avec des pseudo-mots : une revue des corrélations phonosymboliques. *Signifiances (Signifying)*, Le symbole est-il diabolique ? Duplicité(s) du signe en question, 2 (1), 249–268.

Bremner, A. J. *et al.* (2013). "Bouba" and "Kiki" in Namibia? A remote culture makes similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, 126(2), 165–172.

Burton, H., Snyder, A. Z., Diamond, J. B., & Raichle, M. E. (2002). Adaptive changes in early and late blind: a fMRI study of verb generation to heard nouns. *Journal of neurophysiology*, 88(6), 3359–3371.

Campbell, E., Casillas, R., & Bergelson, E. (2024). The role of vision in the acquisition of words: Vocabulary development in blind toddlers. *Developmental Science*, 27, e13475.

Champoux, F., Collignon, O., Bacon, B. A., Lepore, F., Zatorre, R. J., & Théoret, H. (2010). Early- and Late-Onset Blindness Both Curb Audiotactile Integration on the Parchment-Skin Illusion. *Psychological Science*, 22(1), 19-25.

Chastaing, M. (1964). L'opposition des consonnes « sourdes » aux consonnes « sonores » a-t-elle une valeur symbolique ? *Vie et langage*, 147, 367–370.

Correa Silva, M., Bernard, A., & Centre Amandine Talent. (2014). *Images que l'on peut toucher*. Collection Corpus Tactilis, Les Doigts qui rêvent.

Cox, G. E., Hemmer, P., Aue, W. R., & Criss, A. H. (2018). Information and processes underlying semantic and episodic memory across tasks, items, and individuals. *Journal of Experimental Psychology: General*, 147 (4), 545–590.

Cox, G. E., & Criss, A. H. (2020). Similarity leads to correlated processing: A dynamic model of encoding and recognition of episodic associations. *Psychological review*, 127(5), 792–828.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268–294.

Cuskley, C., Simner, J., & Kirby, S. (2017). Phonological and orthographic influences in the bouba-kiki effect. *Psychological research*, 81(1), 119–130.

Ćwiek, A. *et al.* (2021). The bouba/kiki effect is robust across cultures and writing systems. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 377.

Ćwiek, A. *et al.* (2024). The Alveolar Trill Is Perceived as Jagged/Rough by Speakers of Different Languages. *J. Acoust. Soc. Am.* 156, 3468–3479.

- Davidson, P. W. (1972). Haptic judgments of curvature by blind and sighted humans. *Journal of Experimental Psychology*, 93(1), 43–55.
- Davis, R. (1961). The fitness of names to drawings: a cross-cultural study in Tanganyika. *British Journal of Psychology*, 52, 259–268.
- Deroy, O. *et al.* (2016). Differentiated Audio-Tactile Correspondences in Sighted and Blind Individuals. *Journal Of Experimental Psychology. Human Perception and Performance*, 42(8), 1204–1214.
- Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed–accuracy study. *Cognitive Psychology*, 16(4), 519–555.
- Dosher, B. A., & Rosedale, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory and Language*, 30(2), 125–160.
- D’Onofrio, A. (2014). Phonetic Detail and Dimensionality in Sound-shape Correspondences: Refining the Bouba-Kiki Paradigm. *Language and Speech*, 57(3), 367–393.
- Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound-Symbolism is Disrupted in Dyslexia: Implications for the Role of Cross-Modal Abstraction Processes. *Cognitive Science*, 602–607.
- Eitan, Z., Ornoy, E., & Granot, R. Y. (2012). Listening in the dark: Congenital and early blindness and cross-domain mappings in music. *Psychomusicology: Music, Mind, and Brain*, 22(1), 33–45.
- Ernst, M. O. (2005). A Bayesian View on Multimodal Cue Integration. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out: Advances in visual cognition* (pp. 105–131). Oxford University Press.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal Of Vision*, 7(5), 1–14.
- Etzi, R. *et al.* (2016). When Sandpaper Is 'Kiki' and Satin Is 'Bouba': an Exploration of the Associations Between Words, Emotional States, and the Tactile Attributes of Everyday Material. *Multisensory Research*, 29(1–3), 133–155.
- Fort, M., Martin, A., & Peperkamp, S. (2015). Consonants are More Important than Vowels in the Bouba-kiki Effect. *Language and Speech*, 58(2), 247–266.
- Fort, M. *et al.* (2018). Symbouki: a meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental science*, 21(5), e12659.
- Fort, M., & Schwartz, J. L. (2022). Resolving the bouba-kiki effect enigma by rooting iconic sound symbolism in physical properties of round and spiky objects. *Scientific reports*, 12(1), 19172.
- French, P. L. (1977). Towards an explanation of phonetic symbolism. *Word*, 28, 305–322.
- Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influences haptic–auditory associations in the “Bouba–Kiki” effect. *Cognition*, 132(2), 164–173.
- Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & psychophysics*, 68(7), 1191–1203.
- Gordon, I. E., & Morison, V. (1982). The haptic perception of curvature. *Perception & psychophysics*, 31(5), 446–450.
- Gori, M., Sandini, G., Martinoli, C., & Burr, D. (2010). Poor haptic orientation discrimination in nonsighted children may reflect disruption of cross-sensory calibration. *Current Biology*, 20(3), 223–225.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. (2004). Neuropsychology: pitch discrimination in the early blind. *Nature*, 430(6997), 309.

Graven, T., & Desebrock, C. (2018). Bouba or kiki with and without vision: Shape-audio regularities and mental images. *Acta psychologica*, 188, 200–212.

Hamilton-Fletcher, G. *et al.* (2018). The role of visual experience in the emergence of cross-modal correspondences. *Cognition*, 175, 114–121.

Hatwell, Y. (2003). *Psychologie cognitive de la cécité précoce*. Dunod.

Held, R. *et al.* (2011). The newly sighted fail to match seen with felt. *Nature neuroscience*, 14(5), 551–553.

Heller, M. A. *et al.* (2003). Superior haptic perceptual selectivity in late-blind and very-low-vision subjects. *Perception*, 32(4), 499–512.

Heller, M., Gentaz, E. (2018) *Psychologie du toucher et de la cécité*. Talant : Éditions les Doigts qui rêvent.

Hirsch, F. *et al.* (2010) Etude préliminaire de la perception précoce des voyelles labiales par des auditeurs déficients visuels. *Journées d'Etudes sur la Parole*, France.

Hötting, K., & Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychological science*, 15(1), 60–64.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. John Wiley & Sons.

Jakobson, R., (1965). À la recherche de l'essence du langage. *Diogenes*, 51, 22–38.

Jakobson, R., & Waugh, L. (1980) *La charpente phonique du langage*. Paris, Éditions de Minuit, coll. « Arguments », 283.

Kac, M. (1966). Can One Hear the Shape of a Drum? *The American Mathematical Monthly*, 73(4), 1–23.

Knoeferle, K. *et al.* (2017) What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Sci Rep*, 7, 5562.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, 114(1), 19–28.

Lederman, S. J. *et al.* (1990). Visual mediation and the haptic recognition of two-dimensional pictures of common objects. *Attention, Perception, & Psychophysics*, 47(1), 54–64.

Lederman, S. J., & Klatzky, R. L. (1997). Relative availability of surface and object properties during early haptic processing. *Journal of Experimental Psychology: Human Perception and Performance*, 23(6), 1680–1707.

Lewkowicz D. J. (2014). Early experience and multisensory perceptual narrowing. *Developmental psychobiology*, 56(2), 292–315.

Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1–36.

Locke, J. (1689). *An essay concerning human understanding*. New York: Oxford University Press.

Masclé, C. (2020). *Le livre tactile illustré : influence de la technique d'illustration sur la reconnaissance des images par des personnes voyantes et non-voyantes*. Thèse de doctorat, Université Toulouse le Mirail.

- Maurer, D., Pathman, T. & Mondloch, C.J. (2006), The shape of boubas: sound–shape correspondences in toddlers and adults. *Developmental Science*, 9 : 316–322.
- McCormick, K. *et al.* (2015). Sound to Meaning Mappings in the Bouba-Kiki Effect. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1565–1570). Austin, TX : *Cognitive Science Society*.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264 (5588), 746-748.
- Ménard, L. *et al.* (2009). Production and perception of French vowels by congenitally blind adults and sighted adults. *The Journal Of the Acoustical Society of America*, 126(3), 1406–1414.
- Ménard, L. *et al.* (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *The Journal of the Acoustical Society of America*, 134(4), 2975–2987.
- Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal Of Experimental Psychology. Learning, Memory, and Cognition*, 38(5), 1152–64.
- Moos, A. *et al.* (2008). Perception of Ultra-Fast Speech by a Blind Listener – Does he use his visual system? In R. Sock, S. Fuchs, & Y. Laprie (Ed.), *8th International Seminar on Speech Production*. Strasbourg, France : INRIA.
- Motamedi, Y. *et al.* (2019). The iconicity toolbox: Empirical Approaches to measuring iconicity. *Language and Cognition: An Interdisciplinary Journal Of Language and Cognitive Science*, 11(2), 188–207.
- Millar, S. (1997). [Ed. 2013]. *Lire par le toucher*. Talant : Éditions les Doigts qui rêvent.
- Newman, S.S. (1933). Further Experiments in Phonetic Symbolism. *American Journal Of Psychology*, 45, 53–57.
- Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal Of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 65(2), 115–124.
- Nielsen, A., & Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word materials and their implications for language learning and transmission. *Language and Cognition*, 4(2), 115–125.
- Nobile, L. (2015). Phonemes as Images. An experimental inquiry into shape-sound symbolism applied to the distinctive features of French. In: M. K. Higara, W. J. Herlofsky, K. Shinohara & K. Akita (dirs), *Iconicity: East Meets West* (pp. 71–91). John Benjamins Publishing Company.
- Norman, J. F., & Bartholomew, A. N. (2011). Blindness enhances tactile acuity and haptic 3-D shape discrimination. *Attention, perception & psychophysics*, 73(7), 2323–2331.
- O'Boyle, M. W., & Tarte, R. D. (1980). Implications for phonetic symbolism: the relationship between pure tones and geometric figures. *Journal Of psycholinguistic research*, 9(6), 535–544.
- Oberman, L. M., & Ramachandran, V. S. (2008). Preliminary evidence for deficits in multisensory integration in autism spectrum disorders: the mirror neuron hypothesis. *Social neuroscience*, 3(3-4), 348–355.
- Ocelli, V., Bruns, P., Zampini, M., & Röder, B. (2012). Audiotactile integration is reduced in congenital blindness in a spatial ventriloquism task. *Neuropsychologia*, 50(1), 36–43.
- Ocelli, V. *et al.* (2013). The Takete–Maluma Phenomenon in Autism Spectrum Disorders. *Perception*, 42(2), 233–241.

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. Ohala (Eds.), *Sound symbolism* (pp. 325–347). Cambridge, UK : *Cambridge University Press*.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal Of Experimental Child Psychology*, 114(2), 173–186.

Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Experimental Brain Research*, 220(3-4), 319–33.

Piller, S., Senna, I., & Ernst, M. O. (2023). Visual Experience shapes the Bouba-Kiki effect and the size-weight illusion upon sight restoration from congenital blindness. *Scientific reports*, 13(1), 11435.

Platon. (ed. 1846) *Cratyle, ou de la propriété des noms. Rey et Gravier*, traduit par Victor Cousin.

Preziosi, M. A., & Coane, J. H. (2017). Remembering that big things sound big: Sound symbolism and associative memory. *Cognitive research: principles and implications*, 2(1), 10.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia: a window into perception, thought and language. *Journal Of Consciousness Studies*, 8, 3–34.

Révész, G. (1950). *Psychology and art of the blind*. Longmans, Green.

Röder, B., Rösler, F., & Neville, H. J. (2000). Event-related potentials during auditory language processing in congenitally blind and sighted people. *Neuropsychologia*, 38(11), 1482–1502.

Sadato, N. *et al.* (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 380(6574), 526–528.

Sadato, N., Okada, T., Honda, M., & Yonekura, Y. (2002). Critical period for cross-modal plasticity in blind humans: a functional MRI study. *NeuroImage*, 16(2), 389–400.

Sapir, E. (1929). A study in phonetic symbolism. *Journal Of Experimental Psychology*, 12, 225–239.

Sathian, K., & Stilla, R. (2010). Cross-modal plasticity of tactile perception in blindness. *Restorative neurology and neuroscience*, 28(2), 271–281.

Saussure, F. de (1916). *Cours de linguistique générale*. Paris: Payot.

Schild, U., & Friedrich, C. K. (2018). What determines the speed of speech recognition? Evidence from congenitally blind adults. *Neuropsychologia*, 112, 116–124.

Schulman, A. I. (1974). Memory for words recently classified. *Memory & Cognition*, 2(1-A), 47–52.

Senna, I. *et al.* (2021). Development of multisensory integration following prolonged early-onset visual deprivation. *Current biology : CB*, 31(21), 4879–4885.

Sidhu, D. M., & Pexman, P. M. (2017). A prime example of the maluma/takete effect? Testing for sound symbolic priming. *Cognitive Science*, 41(7), 1958–1987.

Sidhu, D. M., & Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review*, 25, 1619–1643.

Sidhu, D. M. *et al.* (2021). Sound symbolism shapes the English language: The maluma/takete effect in English nouns. *Psychonomic Bulletin & Review*, 28(4), 1390–1398.

Sonier, R. P. *et al.* (2020). A round Bouba is easier to remember than a curved Kiki: Sound-symbolism can support associative memory. *Psychonomic Bulletin & Review*, 27(4), 776–782.

Sourav, S. *et al.* (2019). A Protracted Sensitive Period Regulates the Development of Cross-Modal Sound–Shape Associations in Humans. *Psychological Science*, 30(10), 1473–1482.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, and Psychophysics*, 73, 971–995.

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebral cortex (New York, N.Y. : 1991)*, 19(5), 1198–1207.

Sumby, W.H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212–215.

Thorndike, E. L. (1945). On Orr's hypotheses concerning the front and back vowels. *British Journal Of Psychology*, 36, 10–14.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, 93(1), 10–19.

Westbury, C. *et al.* (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal Of Memory and Language*, 99, 122–150.

Winter B, Pérez-Sobrinó P, Brown L (2019) The sound of soft alcohol: Crossmodal Associations between interjections and liquor. *PLoS ONE*, 14(8): e0220449.

Winter, B. *et al.* (2022). Trilled /r/ is associated with roughness, linking sound and touch across spoken languages. *Scientific reports*, 12(1), 1035.

Bibliographie pour les outils informatiques

Hartig, F. (2024). DHARMA: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models. R package version 0.4.7, <https://cran.r-project.org/package=DHARMA>.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, 51(1), 195–203.

Posit team, (2025). RStudio: Integrated Development Environment for R.

The MathWorks Inc. (2019). MATLAB, version 9.7.0 (R2019b). Natick, Massachusetts: The MathWorks Inc. <https://www.mathworks.com>

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S (4th ed.)*. Springer. Package available at: <https://cran.r-project.org/package=MASS>

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer. Package available at: <https://ggplot2.tidyverse.org/>

Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. (2025). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, <https://dplyr.tidyverse.org>.

Annexes

Annexe 1 : Profil des participants non-voyants

Code	Âge	Genre	Catégorie socio-économique	Etudes	Langues maternelles	Locuteur du français	Trouble Cognitif	Main dominante	Braille	Degré de vision	Âge perte vue
D01	34	H	Sans-emploi	Bac+4	Français Farsi	Oui	Non	Droite	Oui	Lumière	15
D02	45	H	Artiste	Bac	Français	Oui	Non	Droite	Oui	Nulle	19
D03	44	F	Profession intellectuelle	Bac+3	Français	Oui	Non	Droite	Oui	Nulle	18
D04	29	H	Profession intellectuelle	Bac+3	Arabe tunisien	Oui	Non	Droite	Oui	Lumière	15
D05	31	H	Etudiant	Bac+5	Arabe algérien	Oui	Non	Droite	Oui	Lumière	13
D06	38	H	Employé	Bac+1	Français	Oui	Non	Ambidextre	Oui	Lumière	0
D07	41	F	Profession intellectuelle	Bac+4	Français	Oui	Non	Droite	Oui	Nulle	0
D08	37	F	Sans-emploi	Bac	Français	Oui	Non	Ambidextre	Oui	Nulle	0
D09	41	H	Employé	Bac+3	Français	Oui	Non	Ambidextre	Oui	Lumière	0

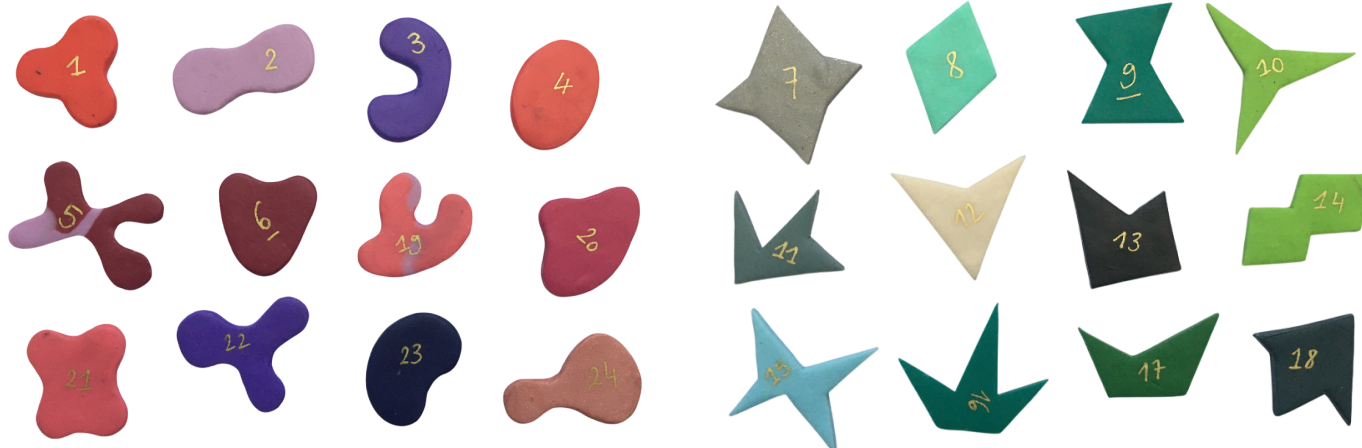
Annexe 2 : Profil des participants voyants

Code	Âge	Genre	Catégorie socio-économique	Etudes	Langues maternelles	Locuteur du français	Trouble Cognitif	Main dominante	Braille
V11	25	H	Artiste	Bac+2	Français	Oui	Non	Droite	Non
V12	31	H	Etudiant	Bac+3	Français	Oui	Non	Droite	Non
V13	48	F	Profession intellectuelle	Bac+5	Français	Oui	Non	Droite	Non
V14	31	H	Employé	Bac	Pashto	Oui	Non	Droite	Non
V15	35	F	Profession intellectuelle	Bac+4	Français	Oui	Non	Droite	Non
V16	52	F	Artiste	Bac+3	Français	Oui	Non	Droite	Non
V17	21	F	Etudiant	Bac+3	Français	Oui	Non	Gauche	Non
V18	50	H	Artiste	Bac+3	Français	Oui	Non	Droite	Non
V19	40	H	Profession intellectuelle	Bac+5	Français	Oui	Non	Droite	Non
V20	23	F	Etudiant	Bac+3	Français Espagnol	Oui	Non	Droite	Non
V21	45	F	Employé	Brevet	Français, créole capverdien	Oui	Non	Droite	Non

Annexe 3 : Paires de forme

LES 24 PAIRES DE FORME

Numéro paire	STIMULUS SONORE	Nature phonombolique du stimulus sonore	Condition de congruence	STIMULUS TACTILE (corrélat haptique du logatome de forme)
F1	monulo	{rond}	congruent	rond
F2	numupo	{rond}	congruent	rond
F3	nojomu	{rond}	congruent	rond
F4	juwolu	{rond}	congruent	rond
F5	wulojo	{rond}	congruent	rond
F6	lonuwu	{rond}	congruent	rond
F7	sikefi	{angulaire}	congruent	angulaire
F8	fefeti	{angulaire}	congruent	angulaire
F9	fepife	{angulaire}	congruent	angulaire
F10	piteke	{angulaire}	congruent	angulaire
F11	tisife	{angulaire}	congruent	angulaire
F12	kepisi	{angulaire}	congruent	angulaire
F13	mupono	{rond}	incongruent	angulaire
F14	nojuwo	{rond}	incongruent	angulaire
F15	numulo	{rond}	incongruent	angulaire
F16	jonowu	{rond}	incongruent	angulaire
F17	woluju	{rond}	incongruent	angulaire
F18	lupomu	{rond}	incongruent	angulaire
F19	sekifi	{angulaire}	incongruent	rond
F20	fitefe	{angulaire}	incongruent	rond
F21	fikipe	{angulaire}	incongruent	rond
F22	pesife	{angulaire}	incongruent	rond
F23	tepefi	{angulaire}	incongruent	rond
F24	kitesi	{angulaire}	incongruent	rond



Annexe 4 : Paires de texture

LES 24 PAIRES DE TEXTURE

Numéro paire	STIMULUS SONORE	Nature phonobolique du stimulus sonore	Condition de congruence	STIMULUS TACTILE (corrélat haptique du logatome de texture)
T1	mowunu	{lisse}	congruent	lisse (coton)
T2	nolupo	{lisse}	congruent	lisse (tissu épais)
T3	nunowu	{lisse}	congruent	lisse (tissu fin_1)
T4	jumulo	{lisse}	congruent	lisse (collant)
T5	wopoju	{lisse}	congruent	lisse (plastique)
T6	lujomo	{lisse}	congruent	lisse (papier)
T7	fikese	{rugueux}	congruent	rugueux (grains semoule)
T8	sepeti	{rugueux}	congruent	rugueux (quadrillage fil de fer)
T9	fivɛfi	{rugueux}	congruent	rugueux (grains de riz)
T10	pefike	{rugueux}	congruent	rugueux (écorce)
T11	ɛɪfipe	{rugueux}	congruent	rugueux (papier de verre)
T12	kesiwi	{rugueux}	congruent	rugueux (éponge verte)
T13	mulowo	{lisse}	incongruent	rugueux (grains de blé concassé)
T14	numujo	{lisse}	incongruent	rugueux (grillage à poule)
T15	poluwo	{lisse}	incongruent	rugueux (plastique troué)
T16	joɲonu	{lisse}	incongruent	rugueux (aluminium froissé)
T17	wujomu	{lisse}	incongruent	rugueux (gant exfoliant)
T18	lonuju	{lisse}	incongruent	rugueux (éponge grattoire)
T19	fɛɪite	{rugueux}	incongruent	lisse (microfibre)
T20	siʃeki	{rugueux}	incongruent	lisse (tissu doux)
T21	ɛɪsipe	{rugueux}	incongruent	lisse (élasthanne)
T22	pefɛwi	{rugueux}	incongruent	lisse (satin)
T23	tikese	{rugueux}	incongruent	lisse (plastique/carton)
T24	ketifi	{rugueux}	incongruent	lisse (tissu fin_2)



Annexe 5 : Questionnaire participant



Questionnaire profil participant-e

Etude psycholinguistique sur l'impact de la vue dans la perception du langage

Vos informations sont personnelles et seront complètement anonymisées pour cette étude. Elles ne seront jamais diffusées en étant associées à votre nom. Les seules personnes susceptibles de retrouver votre identité via le code participant-e sont Lola Corcoral et Anna Marczyk-Bukłaha.

Code participant-e :

Date :

Informations générales :

- 1) Quel âge avez-vous ?
- 2) A quel genre vous identifiez-vous le mieux ?
 - F
 - H
 - Autre identité
- 3) Quel est votre métier ou votre occupation ?
.....
- 4) Avez-vous fait des études ? Si oui, jusqu'à quel niveau ?
.....
- 5) Quelles sont vos langues maternelles ?
.....
- 6) Quelles langues parlez-vous aujourd'hui ?
.....
- 7) Avez-vous une forme d'autisme ? Êtes vous dyslexique ?
 - Oui (si oui, entourer la condition qui vous concerne)
 - Non
- 8) Avez-vous, à votre connaissance, un trouble ou une particularité cognitive ?
 - Oui (si oui, spécifier :))
 - Non
- 9) Est-ce que vous pouvez lire le braille ?
 - Oui
 - Non
- 10) Votre main dominante est la...
 - Droite
 - Gauche
 - Ambidextre

Informations spécifiques au public non-voyant :

11) Comment qualifieriez-vous votre degré de vue actuel ?

- Cécité sévère
- Cécité totale
- Autre

.....

.....

.....

.....

12) A partir de quel âge avez-vous atteint ce degré de vue ?

.....

.....

.....

13) Qu'est-ce qui est à l'origine de votre cécité ?

[QUESTION RETIRÉE DE L'ENTRETIEN POUR MANQUE DE PERTINENCE ET INVASIVITÉ]

Cadre à remplir à la fin de l'expérience.