La langue qui reste

L'acquisition d'une langue par le film, la musique et le dialogue — pourquoi l'ancrage émotionnel tient mieux en mémoire

Dirk Streifler

1re édition · Berlin, 15 octobre 2026

Traduit de l'original allemand ; en cas de doute, la version allemande fait foi.
Ce texte a été traduit de l'original allemand à l'aide de l'IA.
CC BY 4.0 · creativecommons.org/licenses/by/4.0/

Résumé

Ce livre expose le concept pédagogique de Pontivo, une plateforme d'acquisition des langues par le film, le chant et la communauté. Son point de départ : malgré des efforts considérables, un enseignement centré sur les règles produit de façon fiable un savoir sur la langue, mais trop rarement une capacité à s'en servir. Depuis quatre décennies, la recherche nomme les conditions dans lesquelles les êtres humains acquièrent réellement une langue : de grandes quantités d'input compréhensible, avec une anxiété faible.

À partir de ces conditions, le livre élabore la grille d'examen « Compréhensible ? Ancré ? Relié ? » et la promène le long de la plateforme — le film avec sa piste de sous-titres bilingue, le chant comme forme de mémoire, la chorale et la salle de classe comme espaces qui abaissent l'anxiété. Dans le chant du pays d'origine traduit, tous les fils se rejoignent ; le dernier chapitre nomme les limites et la liste de mesures. D'un bout à l'autre, les principes étayés et les paris de conception non mesurés restent séparés. Vérifiable plutôt qu'affirmé.

L'idée est née de la détresse des personnes réfugiées ; elle est construite pour toutes celles et tous ceux qui ont besoin d'une nouvelle langue pour en être — la professionnelle venue de l'étranger et sa famille autant que la classe ordinaire plurilingue. Et elle vise plus loin que le vocabulaire : là où un groupe traduit ensemble les chants de ses membres dans la langue commune et les chante, l'intégration devient pour la première fois audible dans les deux sens. Ce qui tient, ce livre l'étaye ; ce qui reste un pari, il le nomme — et il invite la pédagogie, les médias et les arts à en faire ensemble l'épreuve.

Partie I — Le problème

> „... a gifted person ought to learn English (barring spelling and pronouncing), in 30 hours, French in 30 days, and German in 30 years."
> — Mark Twain, A Tramp Abroad, App. D (1880)

Chapitre 1 · Six cents heures — et pas un mot

1.1 Deux scènes

Amal a trente-quatre ans, elle est arrivée de Syrie en Allemagne il y a trois ans. Elle a fait ce qu'on exigeait d'elle, et davantage : elle a suivi le cours d'intégration général — six cents heures de cours de langue, plus cent heures de cours d'orientation —, elle a rempli des cahiers de vocabulaire, travaillé des fiches, répété des tests blancs. À la fin, elle a réussi le « Deutsch-Test für Zuwanderer » avec le résultat B1. Elle appartient ainsi à la meilleure moitié de sa promotion. Le certificat est chez elle, dans une pochette transparente.

À la caisse du supermarché, la caissière demande, vite et par-dessus l'épaule : „Sammeln Se Punkte?" — vous cumulez des points ? Amal ne comprend pas les mots. Elle ne comprend même pas qu'il s'agissait d'une question. Elle sent la file derrière elle, l'attente, les regards. Elle sourit, hoche la tête, paie et ne dit — rien. Sur le chemin du retour, elle se traduit à elle-même la phrase qu'elle n'a pas dite. Chez elle, elle saurait expliquer sans faute quand l'allemand emploie le datif.

La seconde scène se joue mille kilomètres plus au sud, dans une classe de lycée à Bologne. Ici, l'anglais est à l'emploi du temps depuis l'école primaire ; la plupart de ceux qui sont dans cette salle ont plus de dix ans de cours derrière eux, des programmes de grammaire, des tests de vocabulaire, des devoirs sur table. L'enseignante lance un documentaire en version originale anglaise. Au bout de deux minutes, quelqu'un lève la main et demande s'il n'en existe pas aussi une version italienne. Ce n'est pas une provocation. C'est une information honnête : le son ne les atteint pas.

Les deux scènes sont inventées. Les chiffres qui se tiennent derrière elles ne le sont pas.

1.2 Ce que disent les chiffres

Le cours d'intégration allemand n'est pas une offre accessoire, mais l'instrument central de la politique linguistique du pays : sept cents heures de cours, réglées par la loi, avec l'objectif déclaré d'amener les participantes et les participants au niveau B1 au test final.[1] En 2024, parmi les personnes qui ont passé le test final à l'issue d'un cours d'intégration général, 58,6 pour cent ont atteint ce B1 ; 31,8 pour cent sont restées à A2, 9,6 pour cent en dessous.[2] Il faut lire ce chiffre deux fois. Il ne compte que celles et ceux qui sont effectivement allés jusqu'au test — et même parmi eux, quatre sur dix manquent l'objectif du cours.

Et le chiffre dit encore autre chose, de plus discret. Car B1 n'est pas un niveau élevé. L'échelle de référence européenne le décrit comme un niveau seuil : qui a le B1 peut s'exprimer sur des sujets familiers en phrases simples et reliées entre elles, et comprend les points principaux lorsqu'une langue standard claire est utilisée.[3] La caissière n'utilise pas une langue standard claire. Le quotidien ne le fait presque jamais. Cela signifie : même le fait d'*atteindre* l'objectif du cours ne garantit pas la langue dont on a besoin à la caisse, chez le médecin, à la réunion de parents. Amal n'est pas un cas aberrant du système. Elle en est le cas de réussite.

La scène italienne a sa propre statistique. L'Italie teste ses classes de terminale sur tout le territoire ; en 2025, en anglais, 53,3 pour cent des élèves de la dernière année ont atteint le niveau prévu en compréhension écrite — en compréhension orale, c'était 43 pour cent, et dans le sud du pays nettement moins de trois sur dix.[4] Plus d'une décennie de cours, et moins de la moitié comprend la langue lorsqu'elle est parlée. Qu'on y prenne garde : *comprend*. Parler, il n'en est pas même encore question.

Deux pays, deux systèmes, deux groupes d'apprenants très différents — et le même motif : l'effort est énorme, le rendement en langue vivante ne l'est pas. Celui qui ne connaîtrait qu'une seule des deux scènes pourrait croire à des particularités nationales, à la qualité des cours, à la motivation. Le dédoublement suggère un autre soupçon : cela ne tient pas aux pays. Cela tient à la manière dont nous croyons que les langues entrent dans la tête.

1.3 Un savoir sur la langue, une maîtrise de la langue

Le soupçon porte un nom classique. Au début des années 1980, Stephen Krashen a placé au centre de la recherche sur l'acquisition des langues la distinction dont il est ici question : les adultes disposent de « deux voies distinctes et indépendantes pour développer une compétence dans une langue seconde ».[5] La première voie est l'*apprentissage* — un savoir conscient sur la langue : connaître des règles, nommer des règles, pouvoir parler de règles. La seconde est l'*acquisition* — un processus inconscient, le même par lequel les enfants construisent leur langue maternelle : on ne remarque pas qu'il a lieu, et son résultat n'est pas un savoir de règles, mais une intuition de la langue. Une phrase « sonne juste » ou « sonne faux », sans qu'on puisse dire pourquoi.

Cette distinction explique les deux scènes d'un seul coup. Amal a appris ce que le système de cours promettait d'enseigner : un savoir sur l'allemand. Le datif qu'elle sait expliquer est un savoir véritable — mais c'est un savoir de la mauvaise espèce pour l'instant à la caisse. Là, il n'y a pas le temps d'appliquer des règles ; là, la phrase devrait venir d'elle-même, de l'intuition. Et la classe de Bologne a *appris* l'anglais pendant onze ans en en *acquérant* étonnamment peu — parce que le cours lui a surtout montré des phrases sur la langue, et beaucoup trop rarement la langue elle-même, dans des quantités où une intuition aurait pu naître.

Il faut le dire : la séparation nette des deux voies est restée controversée dans la recherche — que l'appris ne puisse vraiment jamais devenir de l'acquis, comme Krashen l'affirmait, beaucoup en doutent.[6] Pour ce livre, la querelle importe moins qu'il n'y paraît. Car le phénomène vers lequel la distinction pointe n'est pas contesté : le savoir des règles et la maîtrise fluide sont deux choses ; le premier s'installe par l'explication et l'exercice de règles, la seconde, de façon démontrable, pas dans une mesure suffisante. Quatre décennies de recherche n'ont guère changé ce constat de fond.[7]

1.4 Pourquoi cela ne tient pas aux enseignants

Il serait commode de diagnostiquer ici une défaillance : de mauvais cours, des enseignants démotivés, des élèves paresseux. Ce livre ne le fait pas, et ce n'est pas par politesse. Les enseignantes et les enseignants des cours d'intégration et des classes scolaires travaillent dans un système qui leur dicte les conditions : il y a un curriculum qui débouche sur des examens, et les examens interrogent de préférence ce qui peut être interrogé — un savoir de règles, des listes de mots, des formats exercés. Il y a des classes de vingt personnes et plus, où chaque apprenant atteint quelques minutes de parole par heure. Et il y a la peur : celle de la honte en cas d'erreur, celle de l'examen, celle du jugement des autres — une peur qui accompagne chaque heure de cours, précisément chez les adultes qui luttent pour leur place dans le pays d'accueil.

Dans ces conditions, un enseignement centré sur la règle n'est pas une sottise, mais une adaptation rationnelle : il est planifiable, contrôlable, facturable. Seulement, il produit justement ce qu'il contrôle — et non ce dont Amal aurait eu besoin à la caisse. Si l'on ne veut pas accuser le système, mais le comprendre, il faut poser la question autrement : non pas « Qui s'y prend mal ? », mais « Dans quelles conditions la langue naît-elle réellement — et comment produit-on ces conditions, y compris hors de la salle de classe, y compris sans que quiconque doive échouer pour cela ? »

1.5 Ce que ce livre affirme — et ce qu'il n'affirme pas

La thèse de ce livre se dit en trois phrases. La langue s'acquiert, elle ne se bachote pas : elle naît d'un input *compréhensible* — d'une langue que l'on comprend parce que l'image, le contexte et les connaissances préalables aident à la comprendre. Elle reste lorsque cet input est *ancré* — émotionnellement, rythmiquement, mélodiquement, comme reste dans la tête une chanson que l'on n'a jamais cherché à retenir. Et elle se déploie là où les personnes sont *reliées* — là où une communauté abaisse la peur et fait du parler et du chant quelque chose que l'on fait les uns avec les autres au lieu de les uns devant les autres. Compréhensible, ancré, relié : tout ce que proposent les chapitres suivants se mesurera à ces trois conditions.

Pontivo — la plateforme dont ce livre fonde le concept pédagogique — est la tentative de donner à ces trois conditions une forme construite : un film authentique avec des sous-titres bilingues synchrones et une explication au clic pour le passage que l'on vient de ne pas comprendre ; un livre de chants pour le chant commun en groupe ; une salle de classe où un enseignant conduit et où tous peuvent suivre ; un récit qui tient le tout ensemble.[8]

Ce qui est tout aussi important est ce que ce livre *n'affirme pas*. Il n'affirme pas que Pontivo aurait prouvé ce qu'il construit. Les principes sur lesquels la plateforme repose sont étayés par des décennies de recherche ; cela sera démontré chapitre après chapitre, références à l'appui. Mais la forme construite concrète — ce lecteur, ce livre de chants, cette bande dessinée, dans cette combinaison — est une série de paris de conception qui ne sont pas encore mesurés empiriquement. Le livre tiendra les deux plans rigoureusement séparés et dressera, au dernier chapitre, une liste honnête de ce qu'il faudrait mesurer. La norme à laquelle cette collection s'est engagée vaut ici aussi : vérifiable plutôt qu'affirmé — vers l'extérieur comme envers nous-mêmes.

1.6 Le chemin à travers ce livre

Le chemin est organisé en six parties. La partie I nomme le problème et pose, au chapitre 2, le fondement : les conditions dans lesquelles les êtres humains acquièrent effectivement des langues — c'est là que les trois questions « Compréhensible ? Ancré ? Relié ? » recevront leur fondement scientifique. La partie II se tourne vers l'input : vers le film avec ses deux pistes de sous-titres et vers la question de savoir pourquoi l'image en mouvement accompagnée de texte est un input compréhensible d'une densité rare — et où se situent les limites de cette promesse. La partie III appartient à la mélodie : au constat que la langue chantée adhère autrement que la langue parlée, et aux conditions dans lesquelles cela vaut. La partie IV entre dans l'espace social : le chœur, qui relie de façon démontrable plus vite que presque toute autre activité commune, et la salle de classe, où la langue maternelle passe du facteur de perturbation à l'échafaudage. La partie V réunit les trois fils en un seul exercice : la chanson venue du pays d'origine, qu'un groupe traduit ensemble dans la nouvelle langue et chante ensuite ensemble — c'est là qu'apparaîtra pourquoi ce format discret est l'endroit le plus dense de tout le concept, et quel rôle y joue le récit « Coro One Sky ». La partie VI, enfin, revient à Amal et à la classe de Bologne : elle décrit les trois scénarios d'apprentissage concrets pour lesquels Pontivo est construit, trace les limites de l'approche — et fixe la barre à laquelle celle-ci devra se laisser mesurer.

Au commencement, il y avait une femme qui a payé en silence, bien qu'elle possède un certificat. À la fin de ce livre, il doit être compréhensible pourquoi six cents heures de cours n'ont pas pu lui enlever ce silence — et ce qu'il faut à la place. Le chapitre suivant commence là où commence toute acquisition d'une langue : à la question de ce qui se passe dans la tête lorsqu'un être humain n'apprend pas une langue, mais la reçoit.

Notes

1. § 10 al. 1 phrase 1 du règlement allemand sur les cours d'intégration (IntV) : « Der Integrationskurs umfasst 700 Unterrichtsstunden » — le cours d'intégration comprend 700 heures de cours (gesetze-im-internet.de/intv/__10.html, consulté le 2026-07-24) ; sur ces heures, 600 relèvent du cours de langue, § 11 al. 1 phrase 1 IntV (gesetze-im-internet.de/intv/__11.html). L'objectif B1 du cours est formulé expressément par l'Office fédéral de la migration et des réfugiés (BAMF) dans sa statistique officielle : l'objectif est « que les participantes et les participants atteignent au test final DTZ le niveau B1 du Cadre européen commun de référence pour les langues (CECR) » (rapport sur la statistique des cours d'intégration pour l'année 2024, notes relatives aux tableaux DTZ, p. 17 et s. ; notre traduction). ↩
2. BAMF, *Bericht zur Integrationskursgeschäftsstatistik für das Jahr 2024*, p. 19 (tableau « Sprachniveau ‚Deutsch-Test für Zuwanderer' im Jahr 2024 nach Kursarten und Ergebnis ») : cours d'intégration généraux — 160 804 participants (58,6 %) avec B1, 87 449 (31,8 %) avec A2, 26 314 (9,6 %) en dessous de l'A2, sur un total de 274 567 personnes ayant passé le DTZ (PDF, consulté le 2026-07-24 ; références vérifiées sur le texte intégral du PDF). Deux remarques de lecture : pour chaque personne, c'est le résultat le plus élevé qui est compté (note 1 du tableau) ; et les proportions se rapportent aux seules personnes qui ont passé le test — qui abandonne le cours ou ne se présente pas au test n'apparaît pas dans ce taux. ↩
3. Conseil de l'Europe, *Cadre européen commun de référence pour les langues* (2001), échelle globale, niveau B1 (« Threshold »/niveau seuil) ; formulation du texte paraphrasée d'après l'échelle globale et la grille d'auto-évaluation (version française à vérifier sur l'imprimé ; grille en ligne : europass.europa.eu, consulté le 2026-07-24). ↩
4. INVALSI, présentation des résultats « L'inglese nella Prova INVALSI 2025 tra successi e fragilità » (invalsi.it, consulté le 2026-07-24) : dans la dernière année du secondaire (grado 13), 53,3 % atteignent le niveau prévu en reading et 43 % en listening ; dans le Sud (« Sud » 28,6 %, « Isole » 27,3 %), la compréhension orale est dramatiquement inférieure à celle du Nord-Est (plus de 60 %). La base est le Rapporto Nazionale INVALSI 2025 (à vérifier sur le texte intégral). Les niveaux visés par l'épreuve sont B2 (lycées généraux et techniques) et B1+ (filières professionnelles). ↩
5. Krashen, S., *Principles and Practice in Second Language Acquisition*, Pergamon 1982, cité ici d'après l'édition internet autorisée par l'auteur (2009), p. 10 : « adults have two distinct and independent ways of developing competence in a second language » — notre traduction (PDF sur sdkrashen.com, consulté le 2026-07-24 ; pagination selon celle de l'édition internet, à vérifier sur la première édition imprimée). On y trouve aussi la caractérisation de l'acquisition comme processus inconscient doté d'un « sentiment » de la correction (p. 10 s.). ↩
6. La critique s'attaque à la non-falsifiabilité de la séparation stricte et à la thèse supplémentaire de Krashen selon laquelle l'appris ne pourrait jamais passer dans l'acquis

: McLaughlin, B., *Theories of Second-Language Learning*, Londres 1987, chap. 2 (à vérifier sur l'imprimé) ; du point de vue de la théorie de l'acquisition des habiletés, DeKeyser, R., Skill Acquisition Theory, dans : VanPatten/Williams (dir.), *Theories in Second Language Acquisition*, 2e éd., New York 2015, chap. 5 (à vérifier sur l'imprimé). La version plus modérée, aujourd'hui largement partagée, distingue savoir implicite et savoir explicite et tient leur rapport pour perméable de façon limitée (le « débat sur l'interface ») : Ellis, R., *The Study of Second Language Acquisition*, 2e éd., Oxford 2008, chap. 9 (à vérifier sur l'imprimé). ↩

7. Panorama de la recherche : Lightbown, P./Spada, N., *How Languages are Learned*, 4e éd., Oxford 2013, en particulier le chap. 6 sur les limites de l'enseignement explicite de la grammaire (à vérifier sur l'imprimé). Les méta-analyses de Norris/Ortega 2000 montrent également que l'instruction explicite a des effets mesurables, mais surtout limités au savoir explicite (*Language Learning* 50) — avec la réserve qui y est elle-même nommée, à savoir que les instruments de mesure interrogent majoritairement un savoir proche des règles (ibid., à vérifier sur le texte intégral). ↩
8. Note fonctionnelle (description de produit, non résultat de recherche) : les formes construites décrites dans ce livre sont — là où elles ne sont pas expressément signalées comme intention de construction — en service sur la plateforme (pontivo.de, état au 15 octobre 2026). Là où un chapitre traite d'une fonction construite, une note de ce type renvoie à la mise en œuvre en service. ↩

Chapitre 2 · Les conditions de l'acquisition

2.1 Ce qu'un enfant reçoit

Personne n'a jamais expliqué le datif à un enfant de trois ans. Et pourtant, un jour, il se tient dans la cuisine et dit une phrase au datif — non parce qu'il aurait appliqué une règle, mais parce que la phrase *sonne juste* pour lui. L'enfant n'a pas appris la langue comme une matière scolaire. Il l'a reçue : de milliers d'heures de langue qui lui étaient adressées, enchâssées dans des situations qu'il comprenait — à table, en s'habillant, pendant la lecture à voix haute, dans la dispute pour le dernier bonbon. La compréhension est venue d'abord ; la forme est venue après, d'elle-même.

Il faut être prudent avec cette observation. Les adultes ne sont pas de grands enfants ; ils ont moins de temps, d'autres cerveaux, une image de soi achevée et une première langue qui parle toujours en même temps.[1] Mais la recherche sur l'acquisition des langues a maintenu étonnamment stable, au cours des quatre dernières décennies, une affirmation centrale : le mécanisme qui construit la langue chez l'enfant reste disponible aussi pour l'adulte — à condition qu'on lui donne ce que l'enfant reçoit. La question de ce chapitre n'est donc pas de savoir si les adultes peuvent acquérir des langues, mais sous quelles conditions. Deux d'entre elles peuvent être nommées, documentées — et construites.

2.2 Première condition : un input compréhensible

La première condition, Stephen Krashen l'a résumée dans une formule devenue le sigle le plus connu de la discipline : *i + 1*. Si *i* est le niveau actuel d'un apprenant, alors l'acquisition a lieu lorsqu'il *comprend* une langue qui se situe un cran au-dessus. La formulation de Krashen lui-même est plus précise que ne le laisse supposer sa version populaire abrégée : une « condition nécessaire (mais non suffisante) » pour le passage de *i* à *i + 1* serait « que l'acquérant comprenne un input qui contient *i + 1* — 'comprendre' signifiant ici qu'il est tourné vers le sens et non vers la forme du message ».[2] Deux choses tiennent dans cette phrase. Premièrement : il ne suffit pas d'*entendre* une langue ; elle doit être comprise. Deuxièmement : la compréhension ne réussit justement pas par l'analyse de la forme, mais en passant à côté d'elle — à l'aide de tout ce qui est là en plus de la langue : la situation, l'image, le savoir du monde, le visage de celui qui parle. C'est précisément pour cela que l'on peut comprendre une langue que l'on ne *maîtrise* pas encore. Et c'est précisément en cet instant, dit l'hypothèse, que la maîtrise grandit.

Il s'ensuit un troisième point, souvent négligé, et important pour tout le reste de ce livre : il n'y a pas besoin de doser *i + 1*. Krashen lui-même a souligné que l'input n'a pas à être taillé sur

mesure pour la « structure suivante » — s'il y a assez d'input compréhensible, *i + 1* s'y trouve automatiquement.[3] Le programme qui « traite » aujourd'hui le datif et la semaine prochaine le passé composé tente de piloter quelque chose qui s'installe de soi-même dès que la quantité suffit. La conséquence pratique est un renversement de la logique de pénurie habituelle : le but n'est pas la *bonne portion* de langue, mais la *grande quantité* de langue compréhensible.

Ici encore, il faut rester honnête : que l'input compréhensible soit *nécessaire*, personne ne le conteste dans la recherche ; qu'il soit *suffisant*, si. La recherche sur l'interaction a montré que la négociation du sens dans la conversation est ce qui rend l'input vraiment efficace, et la recherche sur l'output, que parler et écrire soi-même a une fonction propre et irremplaçable — qui ne fait que comprendre n'a jamais à *construire* de phrases.[4] Ce livre ne reprend donc pas la thèse maximale de Krashen, mais le noyau consensuel : sans input compréhensible massif, aucune langue ne naît ; tout le reste se construit sur lui. Ce que cela signifie pour les limites d'une plateforme qui livre avant tout de l'input, le chapitre 9 en dressera ouvertement le bilan.

2.3 L'épreuve des faits : le déluge de livres des Fidji

Des hypothèses de cette ampleur ont besoin d'une épreuve hors de la salle de séminaire. L'une des plus belles vient des années 1980/81, d'écoles primaires rurales des Fidji, où l'anglais était langue seconde et l'enseignement strictement structuré. Warwick Elley et Francis Mangubhai ont équipé des classes des niveaux 4 et 5 de 250 livres d'histoires anglais attrayants chacune — un « déluge de livres » — et les ont laissées y travailler vingt à trente minutes par jour : une moitié en lecture commune à voix haute, l'autre en lecture silencieuse individuelle. Les classes témoins ont gardé leur programme habituel fondé sur les structures. Point décisif : les classes à livres n'ont reçu *aucun temps d'anglais supplémentaire* — les mêmes heures, remplies autrement.[5]

Au bout de huit mois, les classes à livres du niveau 4 avaient amélioré leur compréhension écrite de quinze mois, les classes témoins de six et demi ; en compréhension orale et aux tests de structures, le même motif apparaissait. La deuxième année, l'avance a gagné l'écriture et d'autres matières ; à travers sept tests de langue différents, les tailles d'effet se situaient entre 0,76 et 1,18 — en recherche sur l'éducation, ce sont de grandes valeurs. Des enfants qui avaient simplement reçu, en quantité, des histoires compréhensibles et intéressantes ont dépassé des enfants qui recevaient un enseignement systématique, y compris et précisément dans ce que les autres avaient exercé systématiquement : la grammaire.[6]

Une prudence s'impose : les Fidji, c'était un déluge de *livres*, chez des enfants, dans une salle de classe. Que le constat se transpose à des adultes devant un écran n'est pas pour autant prouvé — ce qui porte, c'est le principe : un input riche, compréhensible et porteur de sens bat l'instruction centrée sur la forme, sans que personne ait à investir plus de temps. Savoir si le

film peut jouer ce rôle pour des adultes, et sous quelles conditions, est la question de la partie II — et elle a sa propre recherche.

2.4 Deuxième condition : un filtre bas

La deuxième condition ne concerne pas l'input, mais l'être humain qui le reçoit. Krashen l'a appelée le *filtre affectif* : l'anxiété, le manque de confiance en soi, le manque de motivation agissent comme un diaphragme entre l'input compris et la partie de la tête qui en construit de la langue. Si le filtre est haut, quelqu'un peut comprendre entièrement un message — et rien ne reste pour autant. La conclusion de Krashen est remarquablement pratique : l'enseignant efficace serait celui qui « livre de l'input et le rend compréhensible dans une situation de faible anxiété ».[7]

L'image du filtre est une métaphore, non un mécanisme démontré, et elle a été critiquée pour cela.[8] Mais le phénomène vers lequel elle pointe fait partie des mieux documentés de toute la discipline. Elaine Horwitz et ses collègues ont donné en 1986 à l'anxiété langagière son nom et son instrument de mesure, et l'ont décrite comme un syndrome autonome : « un complexe singulier de perceptions de soi, de croyances, de sentiments et de comportements », qui naît du caractère unique de l'apprentissage des langues — car aucune autre matière scolaire n'attaque l'image de soi et l'expression de soi comme le fait de parler dans une langue que l'on ne maîtrise pas encore.[9] Quatre décennies de recherche avec cette échelle et des échelles apparentées montrent partout la même chose : plus l'anxiété langagière est élevée, plus la performance est faible, et cela à travers les pays, les langues et les groupes d'âge.[10] Quiconque a déjà été interrogé dans une classe et a entendu son propre cœur plus fort que la question de l'enseignante connaît ce mécanisme de l'intérieur.

Pour la construction d'environnements d'apprentissage, ce n'est pas une note marginale, mais une prescription de conception de premier ordre. Il ne suffit pas de fournir un bon input ; il faut construire la situation de telle manière que personne n'ait à avoir peur de le recevoir et d'y répondre. C'est la raison profonde pour laquelle ce livre consacre une partie entière à la communauté : le chœur, où la voix singulière est protégée dans la masse, et le groupe fermé, où les erreurs se produisent entre amis, ne sont pas — la partie IV le montrera — des accessoires sociaux, mais des abaisseurs d'anxiété, et donc, au sens littéral, des conditions d'acquisition.

2.5 La troisième question : qu'est-ce qui reste ?

Compréhensible et sans anxiété — cela suffit-il ? Pas encore tout à fait. Entre la compréhension d'une phrase et sa possession se dresse un troisième obstacle : la rétention. Chacun connaît de la langue qui a été comprise et a pourtant disparu sans trace — la conférence de la semaine dernière, la conversation d'hier. Et chacun connaît le contraire : le jingle publicitaire de

l'enfance, la chanson encore restituable mot pour mot après trente ans, la réplique de film que l'on cite avec la voix de l'acteur. Il existe manifestement des formes d'input qui *adhèrent* — parce qu'elles sont codées plusieurs fois (image et son à la fois), parce qu'elles sont chargées d'émotion, parce que le rythme et la mélodie les portent, parce qu'une histoire leur donne une place.

Ce livre résume cette troisième condition sous le mot *ancré*. Contrairement à « input » et à « filtre affectif », ce n'est pas un terme technique établi, mais une parenthèse collective propre à ce livre pour plusieurs effets étudiés séparément — le double codage de l'image et du mot, la mémoire du chanté, la force du récit —, qui seront documentés un à un dans les parties II, III et V.[11] La grille d'examen qui traversera ce livre est ainsi complète. Elle consiste en trois questions que l'on peut poser à tout environnement d'apprentissage, à une salle de classe aussi bien qu'à une application :

- **Compréhensible ?** — L'apprenant reçoit-il de la langue en quantité, qu'il puisse *comprendre* parce que le contexte, l'image et les connaissances préalables l'y aident ?
- **Ancré ?** — L'input est-il fait de telle sorte qu'il *reste* — codé plusieurs fois, émotionnel, rythmique, raconté ?
- **Relié ?** — Le tout se produit-il dans un espace qui *abaisse* l'anxiété au lieu de la produire ?

L'enseignement centré sur la règle du chapitre 1 répond non aux trois questions : son input est rare et centré sur la forme, sa matière est construite pour être interrogeable et non pour adhérer, et sa situation d'examen est une machine à anxiété. C'est là, en trois phrases, le diagnostic — et en même temps le plan de construction de tout ce qui suit.

2.6 L'attraction : pourquoi les gens reviennent

Compréhensible, ancré, relié — les trois questions décrivent ce qui arrive à l'input. Une quatrième question décide s'il y a de l'input du tout : pourquoi quelqu'un qui est venu librement le lundi devrait-il revenir le mardi ? La psychologie de la motivation a là-dessus une réponse étonnamment stable. Selon la théorie de l'autodétermination de Deci et Ryan, l'engagement humain se porte lui-même lorsque trois besoins fondamentaux sont satisfaits : l'autonomie (je choisis), le sentiment de compétence (je peux) et l'affiliation sociale (j'en fais partie).[12] Que l'on lise ces trois notions à côté du plan de construction de ce livre : le choix du film, c'est l'autonomie ; la ligne en langue maternelle produit un sentiment de compétence dès le premier soir — on *comprend*, là où l'on échouait dans le cours ; et le groupe, dont traite la partie IV, c'est l'affiliation. Le texte à trous d'une application d'apprentissage ne satisfait aucun des trois besoins : il choisit à la place de l'apprenant, il documente surtout des fautes, et il se traite seul.

Là où les trois besoins se rejoignent, l'activité peut basculer dans l'état que Csíkszentmihályi a appelé le *flow* : le faire se porte lui-même, l'effort disparaît du vécu.[13] La différence est visible à l'œil nu dans le quotidien : le devoir est expédié parce que le système le contrôle ; le film est regardé jusqu'au bout parce que l'on veut savoir comment cela continue. Pour une méthode dont la première condition s'appelle *quantité* (2.2), ce n'est pas une note de bas de page psychologique, mais la réponse à la question de la quantité elle-même : la quantité ne naît pas de la discipline, mais de l'attraction.

2.7 Du principe à la construction

Si l'on prend les trois questions au sérieux, la théorie de l'acquisition des langues devient une tâche de construction : où une femme adulte de Gelsenkirchen, un adolescent de Bologne obtiennent-ils *en masse* de l'input compréhensible dans la langue cible — non pas quarante-cinq minutes par semaine, mais chaque soir, librement, avec plaisir ? Comment cet input doit-il être fait pour qu'il ne soit pas seulement compris, mais qu'il reste ? Et comment construit-on autour de lui un espace social qui abaisse l'anxiété au lieu de l'augmenter ?

Les quatre parties suivantes de ce livre parcourent ces questions l'une après l'autre. La première réponse est plus proche que l'enseignement des langues n'a longtemps voulu l'admettre : elle scintille depuis cent ans sur les écrans de cinéma et de télévision, elle est authentique, émotionnelle, disponible en masse et sous-titrée dans presque toutes les langues du monde. Le chapitre suivant traite du film — et de la question de ce que devient une soirée télévision lorsqu'on la construit correctement.

Notes

1. La comparaison avec l'acquisition de la première langue est une analogie aux limites nettes : les apprenants adultes diffèrent par le budget-temps, les connaissances préalables, l'alphabétisation et par l'influence de la première langue ; des niveaux finaux proches de ceux d'un locuteur natif sont l'exception dans l'acquisition adulte. Krashen lui-même ne prétend que ceci, que les adultes peuvent continuer d'utiliser *le même mécanisme d'acquisition*, non qu'ils le portent aussi loin : Krashen, S., *Principles and Practice in Second Language Acquisition*, Pergamon 1982 (édition internet 2009), p. 10 (PDF, consulté le 2026-07-24). ↩
2. Krashen 1982 (note 1), p. 20 s. : « a necessary (but not sufficient) condition to move from stage i to stage i + 1 is that the acquirer understand input that contains i + 1, where ‚understand' means that the acquirer is focussed on the meaning and not the form of the message » (notre traduction dans le texte ; pagination selon l'édition internet). ↩
3. Krashen 1982 (note 1), p. 20 s. : l'input n'aurait pas à viser *i + 1* de façon ciblée ; « if the acquirer understands the input, and there is enough of it, i + 1 will automatically be

provided ». ↩

4. Sur l'hypothèse de l'interaction, Long, M., The role of the linguistic environment in second language acquisition, dans : Ritchie/Bhatia (dir.), *Handbook of Second Language Acquisition*, San Diego 1996, p. 413–468 (à vérifier sur l'imprimé) ; sur l'hypothèse de l'output, Swain, M., Communicative competence: Some roles of comprehensible input and comprehensible output in its development, dans : Gass/Madden (dir.), *Input in Second Language Acquisition*, Rowley 1985, p. 235–253 (à vérifier sur l'imprimé). Sur la critique générale de l'architecture théorique de Krashen, déjà le chapitre 1, note 6. ↩
5. Elley, W./Mangubhai, F., The impact of reading on second language learning, *Reading Research Quarterly* 19 (1983) 53–67 (à vérifier sur le texte intégral) ; dispositif et résultats restitués ici d'après l'exposé détaillé des auteurs eux-mêmes : Elley, W./Cutting, B./Mangubhai, F./Hugo, C., *Lifting Literacy Levels with Story Books*, 1996, ERIC ED416441, p. 4 s. (texte intégral, consulté le 2026-07-24 ; là expressément : « no extra time was spent on English by the Book Flood Groups »). ↩
6. Elley/Cutting/Mangubhai/Hugo (note 5), p. 5 : niveau 4 — 15 mois de progrès en lecture (les deux groupes à livres) contre 6,5 mois (témoin) après huit mois ; niveau 5 — 15 et 9 respectivement contre moins de 3 mois ; la deuxième année, « highly significant differences » aux sept tests de langue, tailles d'effet 0,76–1,18. Pour situer l'ordre de grandeur : selon la convention usuelle (Cohen), des tailles d'effet à partir de 0,8 sont tenues pour grandes. ↩
7. Krashen 1982 (note 1), p. 30–32 ; le résumé des deux conditions p. 32 (« comprehensible … input containing i + 1 … and second, a low or weak affective filter to allow the input ‚in' ») ; la citation sur l'enseignant p. 32 (« The effective language teacher is someone who can provide input and help make it comprehensible in a low anxiety situation », notre traduction). ↩
8. Krashen situe le filtre lui-même « outside the language acquisition device proper » (1982, note 1, p. 31) — c'est une image explicative pour un rapport, non un mécanisme identifié sur le plan neurophysiologique ; le concept remonte à Dulay/Burt (1977) (ibid., p. 31). De façon critique sur sa vérifiabilité, McLaughlin (chapitre 1, note 6). ↩
9. Horwitz, E. K./Horwitz, M. B./Cope, J., Foreign Language Classroom Anxiety, *The Modern Language Journal* 70 (1986) 125–132, ici p. 128 : « we conceive foreign language anxiety as a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process » (notre traduction dans le texte) ; sur la place à part de la discipline, ibid., p. 128 (« Probably no other field of study implicates self-concept and self-expression to the degree that language study does. »), doi:10.1111/j.1540-4781.1986.tb05256.x (références vérifiées sur le texte intégral). On y trouve aussi la Foreign Language Classroom Anxiety Scale (FLCAS), p. 129 s. ↩
10. Résumé de façon méta-analytique entre autres chez Teimouri, Y./Goetze, J./Plonsky, L., Second language anxiety and achievement: A meta-analysis, *Studies in Second Language*

Acquisition 41 (2019) 363–387 (corrélations négatives de force moyenne d'un bout à l'autre ; à vérifier sur le texte intégral) ; de façon semblable Botes/Greiff/Dewaele 2020 sur la FLCAS (à vérifier sur le texte intégral). Il est renoncé à la reprise de coefficients concrets jusqu'à ce que les sources primaires soient vérifiées. ↩

11. « Ancré » est donc une contribution propre de ce livre comme notion *ordonnatrice*, non comme constat : les éléments — théorie du double codage et apprentissage multimédia (partie II), mémoire de la langue chantée (partie III), structure narrative (partie V) — sont des lignes de recherche étrangères, documentées chacune séparément ; n'est nouveau que leur regroupement sous une condition placée à côté de « compréhensible » et de « relié ». Là où ce livre intègre au lieu de documenter, cela est signalé — comme ici. ↩
12. Deci, E. L./Ryan, R. M., *Intrinsic Motivation and Self-Determination in Human Behavior*, New York 1985 (à vérifier sur l'imprimé) ; de façon synthétique Ryan/Deci, Self-determination theory and the facilitation of intrinsic motivation, *American Psychologist* 55 (2000) 68–78 (à vérifier sur le texte intégral). Avertissement sur la transposition : la théorie de l'autodétermination est une théorie générale de la motivation, non une théorie didactique des langues ; l'attribution de ses trois besoins fondamentaux au choix du film, à la ligne en langue maternelle et au groupe est une interprétation de ce livre, non un résultat de mesure. ↩
13. Csíkszentmihályi, M., *Flow: The Psychology of Optimal Experience*, New York 1990 (à vérifier sur l'imprimé). Ici encore l'avertissement sur la transposition : la recherche sur le flow a été développée sur des activités actives (escalade, pratique musicale, échecs) ; que le visionnage compréhensif avec double piste produise l'ajustement entre exigence et capacité qui y est décrit est plausible, mais non mesuré — cela appartient matériellement à la liste de mesure du chapitre 9 (mesures de rétention et d'usage). ↩

Partie II — L'input : voir, entendre, lire

> „maiorem habere vim ad discenda ista liberam curiositatem quam meticulosam necessitatem."
> — Augustinus, Confessiones I,14,23

Chapitre 3 · Voir, entendre, lire : le film à double codage

3.1 Une expérience européenne due au hasard

Pourquoi les Néerlandais, les Scandinaves, les Portugais parlent-ils un anglais si remarquablement bon — et les Allemands, les Espagnols, les Italiens un anglais tellement moins bon ? La réponse habituelle montre les écoles. La réponse plus intéressante montre le téléviseur. Lorsque, dans les années 1920 et 30, le cinéma parlant est arrivé en Europe, chaque pays a dû décider ce qu'il faisait des films en langue étrangère. Les grands espaces linguistiques qui pouvaient s'offrir le doublage, coûteux — l'Allemagne, l'Italie, l'Espagne, la France —, ont fait réenregistrer les voix. Les petits pays ont sous-titré. La décision est tombée pour des raisons économiques et politiques, bien avant que quiconque pense à la didactique des langues — et elle est restée, comme restent les systèmes médiatiques, pour des générations.

L'histoire a ainsi mis en place une expérience qu'aucun chercheur n'aurait jamais fait autoriser : des populations entières, sur des décennies, l'un des groupes avec le son original anglais et des sous-titres, l'autre avec une version doublée. Une étude économétrique en a dressé le bilan : les pays de tradition sous-titrée obtiennent des résultats nettement meilleurs aux tests d'anglais — l'effet sous-titres estimé correspond à environ 17 pour cent du niveau de compétence moyen, une fois et demie l'écart-type dans la comparaison entre pays, et il est le plus fort exactement là où le téléviseur livre le plus : en compréhension orale.[1] Le Portugal sous-titre, l'Espagne double ; les résultats portugais au TOEFL se situent nettement au-dessus des résultats espagnols — pour deux pays voisins aux langues étroitement apparentées.[2] Il faut lire ces chiffres avec la prudence requise : ce sont des moyennes par pays, non des mesures sur des apprenants individuels, et les pays diffèrent par plus que leurs habitudes télévisuelles.[3] Mais la direction est difficile à écarter par la discussion — d'autant qu'elle pointe exactement là où pointent aussi la recherche de laboratoire et la recherche en salle de classe dont traite ce chapitre.

La soirée télévision, autrement dit, n'a jamais été seulement du divertissement. C'est la plus grande machine à apprendre les langues qui ait jamais été construite — sauf que personne ne l'a construite comme telle. Ce chapitre demande *pourquoi* elle fonctionne, et le suivant, comment on la fait passer du produit du hasard à l'instrument.

3.2 Deux canaux sont plus qu'un

La première raison tient à l'architecture de la mémoire. Allan Paivio a développé, dans les années 1970 et 80, la *théorie du double codage* : les êtres humains traitent l'information dans deux systèmes séparés mais reliés — un système verbal pour les mots et un système imagé pour la représentation. Ce qui est déposé dans les deux systèmes à la fois, un mot *avec* son image, a deux traces mnésiques au lieu d'une et, en conséquence, davantage de chemins pour être retrouvé.[4] Richard Mayer en a construit une théorie de l'apprentissage multimédia largement éprouvée empiriquement, dont le constat de base est entre-temps un savoir de manuel : les êtres humains apprennent mieux du mot *et* de l'image que du mot seul.[5]

Pour l'acquisition d'une langue, c'est plus qu'une technique de mémorisation. Qu'on se souvienne du paradoxe de Krashen, au chapitre 2 : l'acquisition a lieu lorsqu'on comprend une langue que l'on ne maîtrise à vrai dire pas encore — et cela ne réussit que parce qu'il y a, en plus de la langue, quelque chose qui aide à comprendre. Le film est exactement cet « en plus de la langue » : l'action montre de quoi il est question ; le visage montre comment c'est entendu ; la scène livre le contexte dont le débutant a besoin pour faire, d'un flux sonore incompréhensible, un message compréhensible. Le film est *i + 1* avec aide à la compréhension intégrée — de l'input compréhensible en quantité industrielle.

3.3 La troisième piste : des sous-titres dans la langue cible

Mais entendre et voir, à eux seuls, laissent une lacune. Celui qui n'entend qu'une langue étrangère ne sait souvent même pas *où un mot s'arrête et où le suivant commence* — le flux sonore ne connaît pas d'espaces. C'est ici que la troisième piste entre en jeu : le sous-titre dans la langue cible, la lecture simultanée de ce que l'on entend. Dès les années 1980, avec les sous-titres télétexte de la BBC, Robert Vanderplank a observé systématiquement que les apprenants tirent de la langue d'une télévision sous-titrée — et que le sous-titre ne détourne pas de l'écoute, mais rend lisible ce qui est entendu.[6] Depuis, cela est devenu une direction de recherche à part entière. Sa méta-analyse — dix-huit études sur la compréhension orale et le vocabulaire — aboutit à un résultat clair : les sous-titres dans la langue cible améliorent les deux, la compréhension et l'apprentissage des mots, avec de grands effets.[7]

Une étude du Michigan souvent citée montre particulièrement bien pourquoi. Des apprenants d'arabe, de chinois, d'espagnol et de russe ont regardé de courtes vidéos avec et sans sous-titres dans la langue cible ; avec sous-titres, ils ont obtenu de meilleurs résultats aux tests de vocabulaire et de compréhension. L'explication des autrices est la décisive : les sous-titres aident à *identifier les frontières de mots* — ils décomposent le « flux de parole autrement incompréhensible » en unités auxquelles du sens peut s'accrocher, et relient l'entendu à l'écrit.[8] Et que ce mécanisme ne dépende pas de l'enseignement, la recherche sur les enfants le

montre, qui ont simplement *regardé* une télévision en langue étrangère et y ont absorbé de la langue de façon mesurable — incidemment, sans consigne, sans test derrière eux.[9]

3.4 La question de la redondance — tranchée honnêtement

Ce livre doit ici trancher une tension que l'on chercherait en vain dans des textes publicitaires. Car la même recherche sur l'apprentissage multimédia sur laquelle s'appuie la section 3.2 connaît un *principe de redondance* qui semble parler contre le texte à l'écran : qui entend une explication et lit en plus le même texte écrit apprend souvent *moins bien* que sans le texte — la double présentation du même énoncé engorge le traitement au lieu de le soutenir.[10] Si cela était vrai sans restriction, le film sous-titré serait une mauvaise idée.

La résolution tient à l'objectif d'apprentissage, et elle mérite d'être formulée avec précision. Le principe de redondance a été obtenu sur des locuteurs natifs qui devaient apprendre des *contenus* à partir d'explications orales — la formation de l'éclair, les systèmes de freinage. Pour eux, le texte qui défile est effectivement redondant : il ne contient rien que l'oreille ne livre sans peine. Pour l'apprenant de langue, la situation est inverse. Son objectif d'apprentissage n'est pas le contenu derrière la langue, mais la langue elle-même — et son oreille, justement, ne livre *pas* sans peine : elle perd les frontières de mots, avale les terminaisons, devine un mot sur trois. Le sous-titre dans la langue cible n'est pas pour lui une seconde copie du compris, mais souvent le seul accès au non-compris. La littérature sur la redondance nomme elle-même désormais ces conditions limites de façon expresse : chez les non-natifs, avec un son difficilement compréhensible, avec un texte court et présenté lentement, le principe s'inverse ou disparaît.[11] Les constats sur le captioning exposés en 3.3 ne sont donc pas une anomalie qu'il faudrait taire — ils sont le cas particulier bien documenté dans lequel la redondance devient un soutien. De ce débat doit néanmoins rester une épine : le texte à l'écran n'est pas gratuit. Il coûte de l'attention, et un environnement d'apprentissage qui emploie des sous-titres doit accepter qu'on lui demande s'il investit bien cette attention.

3.5 Deux langues à la fois : le pari du bilingue

Pontivo va au-delà du sous-titre monolingue : son lecteur affiche *deux* pistes à la fois — la langue cible, synchrone avec le son, et en dessous la langue maternelle de l'apprenant. La logique découle de tout ce qui précède : la piste en langue cible segmente le flux sonore et relie le son à l'écrit ; la piste en langue maternelle garantit la compréhension — elle maintient la condition « compréhensible » là même où l'image et le contexte ne suffisent plus seuls, et rend par là seulement accessible aux débutants un matériau authentique et non gradué. Elle est, pour anticiper une notion du chapitre 7, du *translanguaging* vécu : la langue maternelle n'est pas mise à la porte, mais attelée comme échafaudage de compréhension.[12]

Que dit la recherche des doubles sous-titres bilingues ? Elle est jeune, restreinte — et encourageante. Une étude oculométrique auprès d'apprenants chinois d'anglais a comparé les quatre conditions (langue maternelle seule, langue cible seule, les deux, aucune) : la double piste n'a produit *aucune* surcharge mesurable par rapport à une piste unique, et elle a aidé la compréhension par rapport au visionnage sans sous-titres ; les regards des apprenants ont réparti le travail d'eux-mêmes — stables sur la ligne en langue maternelle pour le sens, flexibles sur la ligne en langue cible pour la forme.[13] D'autres études trouvent des gains de vocabulaire à partir d'un visionnage doublement sous-titré.[14] Mais il serait malhonnête d'en faire plus que ce que c'est : ce sont peu d'études, avec de petits échantillons et de courts extraits. Que la forme construite concrète de Pontivo — la double piste synchrone en service continu, sur des mois, auprès de personnes réfugiées et de classes scolaires — tienne ce que les principes promettent est un *pari de conception* : bien fondé, pas encore mesuré. Il figure tout en haut de la liste de mesure du chapitre 9.

3.6 Ce que le film ne peut pas encore

Rester honnête veut dire aussi ceci : le film a des limites, et deux d'entre elles sont fondamentales. La première : il livre de la compréhension, mais aucune explication de mot isolé. Celui qui reste accroché à un mot au milieu d'une phrase, la plus belle des doubles pistes ne lui sert à rien — il a besoin d'une réponse sur *ce mot-là*, maintenant, sans quitter le film. Comment, de cet instant du besoin, naît du vocabulaire est la question du chapitre suivant. La seconde limite est plus profonde : le film ne fait de personne un locuteur. Il remplit l'oreille, non la bouche. Ce qu'il faut pour que de la langue comprise devienne de la langue parlée — et pourquoi ce chemin peut passer par le chant —, les parties III et IV en traitent.

Notes

1. Rupérez Micola, A./Aparicio Fenoll, A./Banal-Estañol, A./Bris, A., TV or not TV? The impact of subtitling on English skills, *Journal of Economic Behavior & Organization* 158 (2019) 487–499, doi:10.1016/j.jebo.2018.12.019 (références vérifiées sur le PDF des auteurs) : effet sous-titres « equivalent to 16.9% (one and a half standard deviations) of the average level of English skills » ; le plus fort en compréhension orale. Sur le choix historique des modes de traduction dans les années 1920/30, ibid., section 2. ↩
2. Ibid., section 2 (exemple Portugal/Espagne : « in Spain television is dubbed. And, as a result, Portugal's results in the TOEFL exams are much better than Spain's ») ; en comparaison brute, les pays sous-titreurs devancent les autres de 3,4 points au TOEFL en ligne, le plus nettement en listening (ibid., résultats descriptifs). ↩
3. Avertissement sur la transposition : il s'agit de moyennes par pays (moyennes nationales au TOEFL) avec une identification économétrique par le choix historique du mode de traduction — non de mesures de parcours d'apprentissage sur des individus. Ce qui porte

: du son original sous-titré en masse agit sur les compétences langagières à l'échelle d'une population, en particulier sur la compréhension orale. Ce qui ne porte pas : une affirmation sur un produit déterminé, un groupe d'apprenants déterminé ou la forme du double sous-titre (à ce sujet, note 13 s.). Les questions de sélection (qui se présente au TOEFL ?) sont discutées par les auteurs eux-mêmes (ibid., section 4). ↩

4. Paivio, A., *Mental Representations: A Dual Coding Approach*, Oxford 1986 (à vérifier sur l'imprimé) ; de façon synthétique id., Dual coding theory: Retrospect and current status, *Canadian Journal of Psychology* 45 (1991) 255–287 (à vérifier sur le texte intégral). ↩
5. Mayer, R. E., *Multimedia Learning*, 3e éd., Cambridge 2021 (à vérifier sur l'imprimé ; « principe multimédia »). La théorie de Mayer relie l'hypothèse des deux canaux de Paivio à une capacité de traitement limitée et à un traitement actif ; elle a été développée sur du matériel explicatif (non sur du matériel d'apprentissage des langues) — sur la limite de transposition, voir 3.4 et la note 10 s. ↩
6. Vanderplank, R., The value of teletext sub-titles in language learning, *ELT Journal* 42 (1988) 272–281 (à vérifier sur l'imprimé) ; id., *Captioned Media in Foreign Language Learning and Teaching*, Londres 2016 (à vérifier sur l'imprimé). ↩
7. Montero Perez, M./Van Den Noortgate, W./Desmet, P., Captioned video for L2 listening and vocabulary learning: A meta-analysis, *System* 41 (2013) 720–739, doi:10.1016/j.system.2013.07.013 : grand effet des sous-titres en langue cible sur la compréhension orale (15 études) comme sur le vocabulaire (10 études). Le texte intégral n'était pas librement accessible par voie électronique ; les tailles d'effet rapportées (restituées dans la littérature avec g ≈ 0,99 et 0,87) sont reprises de seconde main et **à vérifier sur le texte intégral** ; n'est repris ici que l'énoncé central du résumé (grand effet sur les deux). ↩
8. Winke, P./Gass, S./Sydorenko, T., The effects of captioning videos used for foreign language listening activities, *Language Learning & Technology* 14/1 (2010) 65–86, ici p. 65 (texte intégral, références vérifiées sur le PDF) : « captioning was more effective than no captioning » ; les sous-titres aident à identifier les frontières de mots et à segmenter le « flux de parole autrement incompréhensible » (ibid., p. 65, notre traduction). Par souci d'honnêteté : dans les entretiens, des apprenants ont aussi rapporté utiliser les sous-titres comme une « béquille » (ibid., p. 65) — la question de la dépendance est réelle et appartient à la liste de mesure (chapitre 9). ↩
9. d'Ydewalle, G./Van de Poel, M., Incidental foreign-language acquisition by children watching subtitled television programs, *Journal of Psycholinguistic Research* 28 (1999) 227–244 (à vérifier sur l'imprimé — les détails du dispositif, en particulier quelle direction de sous-titrage produisait quel gain, ne sont sciemment pas rapportés ici avant que la source primaire ait été vérifiée). ↩
10. Mayer 2021 (note 5), chapitre sur le principe de redondance (à vérifier sur l'imprimé) ; de façon fondamentale sur la « redundancy », également la recherche sur la charge cognitive autour de Sweller/Kalyuga (à vérifier sur l'imprimé). ↩

11. Les conditions limites du principe de redondance — entre autres les non-natifs, une source audio difficilement compréhensible, un texte court et lent — sont expressément traitées comme « boundary conditions » chez Mayer 2021 (note 5) (à vérifier sur l'imprimé). La méta-analyse sur le captioning en L2 (note 7) est le contrepoids empirique du côté de l'apprentissage des langues. ↩
12. Note fonctionnelle : le lecteur à double piste synchrone est en service dans la médiathèque de la plateforme (pontivo.de/mediathek, état au 15 octobre 2026). ↩
13. Liao, S./Kruger, J.-L./Doherty, S., The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension, *The Journal of Specialised Translation* 33 (2020) 70–98, doi:10.26034/cm.jostrans.2020.549 (accès libre, références vérifiées sur le texte intégral) : « Bilingual subtitles … did not appear to induce more cognitive load or produce more cognitive gain than monolingual subtitles » ; par rapport à la condition sans sous-titres, « more beneficial … linguistic support » (résumé, p. 70) ; attention du regard plus stable sur la ligne en L1, ibid. Vingt participants, quatre courtes vidéos — une petite étude, à pondérer comme telle. ↩
14. Wang, A./Pellicer-Sánchez, A., Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study, *Language Learning* 72 (2022), doi:10.1111/lang.12495 (à vérifier sur le texte intégral) ; en outre la recherche ultérieure sur les doubles sous-titres et l'acquisition du vocabulaire (entre autres *ReCALL*, *SSLA* 2022 s., à vérifier sur le texte intégral). La situation d'ensemble : constamment positive, mais jeune et étroite. ↩

Chapitre 4 · Le mot à l'instant du besoin

4.1 L'endroit où l'on reste bloqué

On est dans le film, cela va bien, on comprend — et puis arrive un mot auquel tout est suspendu. *L'essaim se comporte comme un seul animal.* Essaim ? La double piste du chapitre 3 aide jusqu'à un certain point : en dessous se trouve la traduction de la phrase. Mais quel mot dans la phrase était lequel ? Pourquoi le verbe se tient-il là où il se tient ? Et que signifie exactement ce mot unique dont il est manifestement question ? En cet instant, l'apprenant est dans un état que tout didacticien ne peut que souhaiter : il *veut savoir quelque chose*. Non parce que cela figure au programme, mais parce que cela lui barre le film à l'instant même. La question de ce chapitre est ce que l'on fait de ce moment — et ce que deviennent les mots lorsqu'on l'utilise bien.

4.2 L'apprentissage incident est réel — et il est lent

La recherche sur le vocabulaire a une bonne nouvelle et une nouvelle inconfortable. La bonne : le vocabulaire croît effectivement de façon *incidente* — en lisant, en écoutant et en regardant, sans cahier de vocabulaire, simplement parce que des mots reviennent encore et encore dans des contextes compréhensibles. La part prépondérante du vocabulaire que possèdent les locuteurs natifs est née exactement ainsi ; personne ne leur a enseigné vingt mille mots.[1] Devant l'écran aussi, cela fonctionne : dans l'étude de Louvain, qui comparait trois variantes de sous-titrage, les spectateurs avaient absorbé des mots de façon mesurable après deux visionnages de trois extraits français — les groupes avec sous-titres reconnaissaient la *forme* des mots nouveaux significativement mieux que le groupe témoin et savaient encore dans quel extrait ils les avaient rencontrés.[2]

La nouvelle inconfortable se trouve dans la même étude : à la question de ce que les mots *signifient*, l'avance était déjà plus mince — seuls les groupes dont les sous-titres mettaient en relief les mots-clés dépassaient le groupe témoin pour la reconnaissance du sens, et pour le rappel libre du sens l'effet disparaissait entièrement.[3] Ce n'est pas une valeur aberrante, mais le constat normal de l'apprentissage incident : il construit d'abord un sentiment de la forme — *ce mot existe, voilà comment il sonne, voilà où il va* — et seulement au fil de nombreuses rencontres le sens assuré. L'incidence est large, mais plate ; par rencontre, il reste peu, et les mots rares se rencontrent rarement.[4] Qui ne fait que regarder apprend vraiment — mais lentement.

4.3 Ce que change la glose

Il existe un outil éprouvé de longue date qui s'attaque exactement à cette faiblesse : la *glose* — la petite explication de mot en marge du texte, aussi ancienne, dans l'imprimé, que le livre lui-même. La recherche l'a mesurée en profondeur. Une méta-analyse portant sur 42 études et près de quatre mille apprenants aboutit à un résultat remarquablement clair : la lecture avec gloses conduisait à 45,3 pour cent de mots cibles appris au test immédiat (contre 26,6 pour cent sans gloses) et à 33,4 pour cent au test différé (contre 19,8) — la glose double presque l'apprentissage incident.[5] Deux détails de la même analyse sont importants pour tout ce qui suit. Premièrement : les gloses dans la *langue maternelle* des apprenants agissaient mieux que des explications dans la langue cible — une preuve de plus que la langue maternelle est, dans le processus d'apprentissage, un échafaudage et non un adversaire.[6] Deuxièmement : les gloses les plus efficaces étaient celles qui faisaient brièvement *travailler* l'apprenant (par exemple les gloses à choix multiple, où l'on sélectionne le sens qui convient), les plus faibles celles que l'on n'a même pas besoin d'aller chercher.[7]

Le second constat a derrière lui une théorie qu'il faut connaître, parce qu'elle adresse directement l'« instant du besoin » de 4.1 : l'*hypothèse de la charge d'implication* de Laufer et Hulstijn. Elle énonce, en gros, qu'un mot reste d'autant mieux que l'instant d'apprentissage contient davantage de trois ingrédients : le *besoin* (j'ai besoin de ce mot maintenant), la *recherche* (je dois me procurer son sens) et l'*évaluation* (je dois vérifier si le sens convient à mon contexte).[8] Le cahier de vocabulaire du dimanche soir n'a que peu des trois. Le clic dans le film a les trois : le besoin vient de l'action, la recherche est le clic lui-même, et l'évaluation se produit à l'instant où l'explication rencontre la scène que l'on a sous les yeux.

4.4 Le clic — et ce qui, en lui, est un pari

L'explication au clic de Pontivo est la glose traduite dans la forme de son siècle : qui reste bloqué à un passage dans le lecteur l'effleure et reçoit une explication de *ce passage dans ce contexte* — dans sa langue maternelle, sans quitter le film.[9] C'est plus qu'un appel au dictionnaire, car ce qui est expliqué n'est pas le mot dans le vide, mais le passage : la tournure, l'allusion, la grammaire qui a construit la phrase ainsi. D'après tout ce que sait la recherche sur les gloses, c'est exactement la bonne forme de construction : en langue maternelle, mue par le besoin, liée au contexte.[10]

Mais ce livre doit ici poser deux paris ouvertement sur la table. Le premier : les explications sont produites par un modèle de langue — une machine qui produit de très bonnes explications et, à l'occasion, de fausses, et la recherche qui a mesuré les gloses a mesuré des gloses *vérifiées par une rédaction*, non des gloses produites par une machine. Que la glose machinale tienne la qualité sur laquelle repose sa base documentaire est une question ouverte et doit être contrôlée

— le livre frère de cette collection a consacré un ouvrage entier au problème que les renseignements machinaux ont besoin de preuves qui portent.[11] Le second pari concerne la mesure : la recherche sur les gloses sait aussi que l'interruption coûte. Celui qui clique toutes les cinq secondes ne regarde plus un film, mais feuillette un dictionnaire avec des images — et perd exactement ce qui rend le film si précieux comme input : le flux, l'histoire, le fait de rester accroché. Où se situe la juste mesure, entre le clic qui comble un trou et le cliquetis qui démonte la soirée, n'est pas élucidé empiriquement. Les deux figurent sur la liste de mesure du chapitre 9.

4.5 Compris n'est pas encore retenu

Supposons que tout réussisse : le film livre de l'input compréhensible en quantité, la double piste maintient la compréhension à un niveau élevé, le clic approfondit les passages qui coincent. Alors, à la fin de la soirée, l'apprenant a — compris. Mais entre la compréhension et la possession d'une langue il y a la nuit, et la nuit est un tamis. Ce qui est encore là au matin ne se décide pas seulement au nombre des rencontres, mais à la *manière* dont la matière a été déposée dans la tête : à savoir si elle avait une forme à laquelle la mémoire trouve prise. La plus ancienne forme de ce genre de l'humanité n'est pas le tableau ni la règle. C'est le vers, le rythme, la mélodie — la chanson que l'on n'essaie pas de retenir et dont on ne se débarrasse pourtant jamais plus. Pourquoi il en est ainsi, et ce que cela signifie pour l'acquisition d'une langue, est la question de la partie suivante.

Notes

1. De façon fondamentale sur l'apprentissage incident (incidentel) du vocabulaire : Nation, I. S. P., *Learning Vocabulary in Another Language*, 2e éd., Cambridge 2013, en particulier le chap. 4 (à vérifier sur l'imprimé) ; Hulstijn, J., Incidental and intentional learning, dans : Doughty/Long (dir.), *The Handbook of Second Language Acquisition*, Oxford 2003, p. 349–381 (à vérifier sur l'imprimé). Il est renoncé à la restitution de valeurs seuils concrètes (combien de rencontres un mot exige, combien de familles de mots donnent quel taux de couverture d'un texte) jusqu'à ce que les sources primaires aient été vérifiées. ↩
2. Montero Perez, M./Peters, E./Clarebout, G./Desmet, P., Effects of captioning on video comprehension and incidental vocabulary learning, *Language Learning & Technology* 18/1 (2014) 118–141, ici p. 118 (texte intégral, références vérifiées sur le PDF) : 133 étudiants flamands, trois extraits français, vus deux fois ; « the captioning groups scored equally well on form recognition and clip association and significantly outperformed the control group ». ↩
3. Ibid., p. 118 : « Only the keyword captioning and full captioning with highlighted keywords groups outperformed the control group on meaning recognition. Captioning

did not affect comprehension nor meaning recall. » ↩

4. Sur la caractéristique « large, mais plate » de l'apprentissage incidentel, parmi bien d'autres : Nation 2013 (note 1), chap. 4 ; les petits gains par rencontre sont la raison pour laquelle la recherche insiste tant sur la répétition et la fréquence (à vérifier sur l'imprimé). ↩
5. Yanagisawa, A./Webb, S./Uchihara, T., How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis, *Studies in Second Language Acquisition* 42 (2020) 411–438, doi:10.1017/S0272263119000688 (chiffres vérifiés sur le résumé de l'éditeur ; texte intégral à vérifier sur l'imprimé) : 359 tailles d'effet issues de 42 études, N = 3 802 ; mots cibles appris avec gloses 45,3 % (immédiat) et 33,4 % (différé) contre 26,6 % et 19,8 % sans. ↩
6. Ibid. (résumé) : « L1 glosses yielded greater learning than L2 glosses » — sans interaction significative avec le niveau de langue ; la supériorité de la glose en langue maternelle valait donc aussi pour les apprenants avancés. ↩
7. Ibid. (résumé) : gloses à choix multiple les plus efficaces, gloses dans le texte et glossaires les plus faibles. L'interprétation — est efficace ce qui impose un bref traitement actif — est une interprétation des auteurs comme de ce livre, et elle recoupe l'hypothèse de la charge d'implication (note 8). ↩
8. Laufer, B./Hulstijn, J., Incidental vocabulary acquisition in a second language: The construct of task-induced involvement, *Applied Linguistics* 22 (2001) 1–26 (à vérifier sur le texte intégral ; la présentation en trois composantes — need, search, evaluation — suit le construit du titre de l'article et la réception standard, par exemple chez Nation 2013, note 1). ↩
9. Note fonctionnelle : l'explication au clic est en service dans le lecteur de la médiathèque (pontivo.de/mediathek) ; les mots sur lesquels on a cliqué migrent vers un fichier personnel doté d'un mode de répétition et d'un mode d'examen (pontivo.de/review, pontivo.de/pruefung). État au 15 octobre 2026. ↩
10. La conclusion qui va des données sur les gloses (notes 5–7) à l'explication au clic est une transposition aux limites nettes : la méta-analyse porte sur la *lecture* avec des gloses préparées ; le clic porte sur la *vidéo* avec des explications produites. Ce qui porte : en langue maternelle, proche du besoin, actif en traitement. Ce qui ne porte pas : une taille d'effet pour la forme construite concrète de Pontivo — elle est non mesurée (chapitre 9). ↩
11. Sur le problème des renseignements machinaux dont les preuves doivent porter l'énoncé : Streifler, D., *Kalibriertes Vertrauen*, ra.de-Verlag 2026 (volume frère de cette collection, consacré à trust-on.net), là en particulier le chap. 1 sur le *misgrounding*. Pour l'explication au clic de Pontivo, il en découle une dette de qualité : vérification par échantillons des explications produites contre une référence spécialisée, avant que le clic puisse valoir comme « glose au sens de la recherche ». ↩

Partie III — La mélodie

> „Wie wenig gehört zum Glücke! Der Ton eines Dudelsacks. — Ohne Musik wäre das Leben ein Irrthum."
> — Friedrich Nietzsche, Götzen-Dämmerung, Sprüche und Pfeile 33 (1889)

Chapitre 5 · Pourquoi la chanson reste dans la tête

5.1 La mémoire que personne n'a commandée

Presque chacun porte avec lui des vers de chansons dans des langues qu'il n'a jamais apprises. Celui qui était jeune dans les années 1980 peut encore aujourd'hui chanter des refrains italiens sans savoir l'italien ; des enfants reprennent des vers de pop anglaise en en restituant fidèlement les sons, des années avant de *comprendre* le premier mot d'anglais. Personne n'a exercé ces vers, personne n'a voulu les retenir — ils sont simplement restés. Le didacticien des langues Tim Murphey a forgé pour la forme la plus anodine de ce phénomène un beau nom : le phénomène *song-stuck-in-my-head*, la chanson qui continue de tourner dans la tête, qu'on le veuille ou non.[1]

Pour un livre sur l'acquisition des langues, c'est plus qu'une curiosité. Le chapitre 4 s'achevait sur la question de la forme que l'input doit avoir pour survivre à la nuit. Une réponse traîne ici à ciel ouvert depuis des millénaires : avant toute écriture, les êtres humains ont transmis leur savoir en vers, en rythmes et en mélodies — non par plaisir esthétique, mais parce que la parole liée était la seule à passer de tête en tête de façon fiable. La question est de savoir si ce vieux constat résiste à un examen moderne : la langue chantée adhère-t-elle vraiment mieux que la langue parlée — et si oui, sous quelles conditions ?

5.2 Ce que dit le laboratoire : la mélodie décompose le flux

La première réponse moderne vient d'un laboratoire de Marseille et se rattache immédiatement au problème de segmentation du chapitre 3 : où, dans un flux sonore étranger, un mot s'arrête-t-il, où le suivant commence-t-il ? Daniele Schön et ses collègues ont fait écouter à des adultes un flux de parole artificiel — des syllabes enchaînées sans pauses, dont on ne peut extraire des « mots » que par des régularités statistiques. Lorsque le flux était présenté *parlé*, les auditeurs restaient au niveau du hasard après sept minutes. Lorsque le même flux était *chanté* — chaque syllabe couplée de façon conséquente à une hauteur de son —, ils reconnaissaient les frontières de mots.[2] La conclusion des auteurs : c'est précisément dans la phase initiale, lors de la première décomposition d'une langue nouvelle, que « les propriétés motivationnelles et structurantes de la musique » agissent ensemble — la mélodie ne rend pas seulement le flux plus agréable, elle l'*articule*.[3]

Le deuxième élément vient d'Édimbourg et concerne non l'écoute, mais la parole. Karen Ludke, Fernanda Ferreira et Katie Overy ont fait apprendre à soixante adultes, pendant quinze

minutes, des phrases hongroises du quotidien — selon le procédé écouter-et-répéter, en trois groupes : parler, parler en rythme, chanter. Le groupe qui chantait a obtenu globalement les meilleurs résultats, et son avance était statistiquement significative précisément aux deux tests qui pèsent le plus lourd : le *rappel et la production* des phrases hongroises parlées.[4] Ce qui est remarquable, c'est ce qui n'expliquait *pas* la différence : ni l'âge ni l'humeur, ni la mémoire de travail phonologique, ni le don ou la formation musicale.[5] Il n'était pas nécessaire d'être musicien — le chant lui-même faisait le travail. Les auteurs résument leur interprétation dans une formule que ce livre reprend : une « méthode écouter-et-chanter » peut soutenir la mémoire littérale des phrases en langue étrangère.[6]

Pourquoi cela est plausible peut se reconstituer à partir des chapitres précédents : la chanson est un input codé plusieurs fois au sens du chapitre 3 — texte, mélodie, rythme, souvent encore mouvement —, elle impose la répétition sans lasser (personne ne lit un texte vingt fois, mais chacun chante un refrain vingt fois), elle ralentit et étire les syllabes, rend donc la figure sonore plus audible qu'une parole rapide, et elle accroche la langue à l'émotion. Mais la plausibilité est bon marché. La vérité est plus compliquée, et elle appartient à la ligne principale, non aux petits caractères.

5.3 Honnêtement : quand la chanson n'aide pas

Car la même recherche connaît des constats qui pointent dans la direction inverse. Isabelle Peretz — l'une des principales chercheuses en cognition musicale — a fait apprendre à des adultes, avec Amélie Racette, des vers de chansons et a trouvé ceci : en chantant, *moins* de mots étaient rappelés qu'en parlant, et que le matériau fût présenté chanté ou parlé ne faisait aucune différence pour la mémoire du texte.[7] Et Wanda Wallace avait montré dès 1994 que la mélodie ne porte le texte que si elle *revient* d'une strophe à l'autre et qu'elle est facile à apprendre ; une mélodie entendue une seule fois ou constamment changeante n'aide pas — elle coûte.[8]

Cela réfute-t-il la partie III de ce livre ? Non — cela la *précise*, et précisément là où le livre de chants de Pontivo est construit. Si l'on lit les constats côte à côte, ils ne donnent pas une cacophonie, mais une liste de conditions. La chanson aide lorsque la mélodie est *simple* et qu'elle se *répète* — alors elle prend en charge le travail d'articulation que décrit Wallace, au lieu d'être elle-même matière à apprendre ; là où la mélodie est difficile, elle devient une double tâche, et le texte perd (c'est le constat de Racette : ses sujets devaient apprendre *en plus* des mélodies inconnues). La chanson aide en outre à retenir *littéralement* des tournures et des phrases — les tests de production de Ludke —, non comme substitut à la compréhension large que livre le film. Et elle aide davantage celui qui *chante avec* que celui qui écoute seulement : chanter est un exercice de prononciation avec métronome intégré.[9] Bref : la chanson n'est pas

un remède miracle, mais un outil spécialisé — l'outil d'ancrage pour les tournures qui doivent tenir.

5.4 Du laboratoire au livre de chants

C'est exactement comme cet outil spécialisé que le livre de chants est construit dans Pontivo. Il mise sur des chansons simples et riches en répétitions — chansons populaires, chansons d'enfants, canons, chansons modernes à chanter ensemble —, donc sur le répertoire qui remplit de lui-même la condition de Wallace : des mélodies de strophe qui se répètent, des refrains qui reviennent. Il affiche le texte et les accords ensemble (au format ChordPro, le standard de la pratique du chant en commun), pour qu'une guitare ou un piano puisse porter le groupe. Et il n'est — c'est le point décisif, celui qui conduit à la partie suivante — pas un automate à karaoké pour individus, mais un *outil de groupe* : le mode enseignant et élève laisse une personne choisir la chanson, et tous chantent la même, au même moment, dans la même pièce.[10]

Ici aussi vaut la règle d'honnêteté de ce livre : que *ce* livre de chants, avec *ces* chansons, dans *cet* encastrement, traduise l'avance de laboratoire du chant en acquisition mesurable d'une langue est un pari de conception — les études de laboratoire travaillaient avec des tâches de quelques minutes et des individus, non avec des mois et des classes.[11] Mais le pari est placé avec précision : il mise sur le cas le mieux documenté (mélodies simples et répétées ; rétention littérale ; chant actif) et évite le cas réfuté (mélodies difficiles, écoute passive, chanson comme substitut universel à l'input).

Une chose cependant, la chanson la peut, qu'aucune étude de laboratoire ne mesure et qu'aucun film ne remplace : on peut la faire *ensemble*. Une classe ne peut pas dire ensemble une phrase à la caisse du supermarché — mais elle peut chanter ensemble, dès le premier jour, y compris celles et ceux qui ne parlent pas encore un mot. Ce qui arrive alors à un groupe d'êtres humains — à son anxiété, à son lien, à son sentiment d'appartenance — est peut-être le constat isolé le plus fort de tout ce livre. C'est le thème du chapitre suivant.

Notes

1. Murphey, T., The song stuck in my head phenomenon: A melodic din in the LAD?, *System* 18 (1990) 53–64 (à vérifier sur l'imprimé) ; id., *Music and Song* (Oxford 1992, à vérifier sur l'imprimé). L'observation centrale de Murphey — les chansons repassent involontairement dans la tête et livrent ainsi de la répétition gratuite — n'est reprise ici que comme phénomène nommé, non comme preuve d'efficacité. ↩
2. Schön, D./Boyer, M./Moreno, S./Besson, M./Peretz, I./Kolinsky, R., Songs as an aid for language acquisition, *Cognition* 106 (2008) 975–983, doi:10.1016/j.cognition.2007.03.005 (constat central vérifié sur le résumé/PubMed ;

détails des trois expériences à vérifier sur le texte intégral) : « a strong learning facilitation of song compared to speech ». La version en sept minutes du constat (flux de syllabes chantés → frontières de mots reconnues, version parlée monotone → niveau du hasard) est restituée ici d'après l'exposé de Ludke et al. 2014 (note 4, p. 41 s.). ↩

3. Schön et al. 2008 (note 2), résumé : l'acquisition initiale « may largely benefit of the motivational and structuring properties of music in song » (notre traduction). ↩
4. Ludke, K. M./Ferreira, F./Overy, K., Singing can facilitate foreign language learning, *Memory & Cognition* 42 (2014) 41–52, ici p. 41, doi:10.3758/s13421-013-0342-5 (références vérifiées sur le PDF des auteurs) : 60 adultes, conditions parler/parler en rythme/chanter, 15 minutes de hongrois ; « superior overall performance » du groupe qui chantait, « statistically significant ($p < .05$) for the two tests that required participants to recall and produce spoken Hungarian phrases ». ↩
5. Ibid., p. 41 : « The differences in performance were not explained by potentially influencing factors such as age, gender, mood, phonological working memory ability, or musical ability and training. » ↩
6. Ibid., p. 41 : « a ‚listen-and-sing' learning method can facilitate verbatim memory for spoken foreign language phrases » (notre traduction dans le texte). ↩
7. Racette, A./Peretz, I., Learning lyrics: To sing or not to sing?, *Memory & Cognition* 35 (2007) 242–253 (constat central vérifié sur le résumé PubMed, texte intégral à vérifier sur l'imprimé) : « Fewer words were recalled when singing than when speaking » ; la forme de présentation (chantée/parlée) n'avait aucune influence sur la mémoire du texte. Important pour situer le résultat : ce qui était appris, c'étaient des chansons non familières — le texte *et* la mélodie étaient nouveaux, le chant était donc une double tâche. ↩
8. Wallace, W. T., Memory for music: Effect of melody on recall of text, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20 (1994) 1471–1485 (constat restitué d'après le résumé et la réception standard, à vérifier sur le texte intégral) : la mélodie soutient le rappel du texte lorsqu'elle revient d'une strophe à l'autre et qu'elle est facile à apprendre ; avec une mélodie entendue une seule fois ou changeante, aucun avantage, voire un désavantage. L'interprétation de Wallace — la mélodie « segmente » le texte, marque les longueurs de vers et les schémas d'accentuation — est l'explication structurelle du constat de Schön (note 2) à l'échelle de la chanson. ↩
9. L'interprétation par la prononciation (le chant étire les syllabes, impose l'articulation dans la mesure) est une interprétation de ce livre, sur la base du constat du test de production de Ludke (note 4) et de la littérature sur la prosodie (parmi bien d'autres : Patel, A., *Music, Language, and the Brain*, Oxford 2008, à vérifier sur l'imprimé) ; comme preuve d'efficacité autonome d'un entraînement de la prononciation par le chant, elle resterait à vérifier. ↩
10. Note fonctionnelle : le livre de chants est en service (pontivo.de/liederbuch, état au 15 octobre 2026) : textes avec accords de guitare (ChordPro), transposition, vue scène en

grands caractères, chant commun en groupe — utilisable sans compte. ↩

11. Avertissement sur la transposition : d'un apprentissage de laboratoire de 15 minutes (note 4) et de flux de langue artificielle (note 2) à un chant de groupe sur des mois, aucune conclusion directe ne mène ; ce qui porte, ce sont le mécanisme (segmentation, répétition, rétention littérale) et la liste de conditions (simple, répété, actif). L'effet de terrain du livre de chants appartient à la liste de mesure (chapitre 9). ↩

Partie IV — La communauté

> „Wo man singet, laß dich ruhig nieder, / Ohne Furcht, was man im Lande glaubt; / Wo man singet, wird kein Mensch beraubt; / Bösewichter haben keine Lieder."
> — Johann Gottfried Seume, Die Gesänge (1804)

Chapitre 6 · Le chœur : l'appartenance abaisse l'obstacle

6.1 Quatre-vingt-dix minutes

Imaginons un mardi soir dans une salle paroissiale. Quatorze personnes qui ne se connaissent pas, deux d'entre elles dans le pays depuis quelques mois seulement. À sept heures, elles se tiennent, gênées, entre les rangées de chaises. À huit heures et demie, elles se tiennent en demi-cercle, ont raté trois fois le même canon et ri deux fois, et lorsque quelqu'un propose de commencer dix minutes plus tôt la semaine suivante, tous acquiescent. Quelque chose s'est passé en ces quatre-vingt-dix minutes qui, dans les cours de langue, ne se passe souvent pas en des mois.

Ce quelque chose est mesurable. Une équipe de chercheurs autour du psychologue de l'évolution Robin Dunbar a suivi pendant sept mois des cours du soir pour adultes nouvellement créés — des classes de chant et, à titre de comparaison, des cours de bricolage et d'écriture — et a fait évaluer aux participants, à trois moments, à quel point ils se sentaient proches de leur groupe. Le résultat : après le premier mois, les chanteurs montraient un gain de proximité ressentie significativement plus grand que tous les groupes de comparaison ; les autres n'ont rattrapé qu'après des mois.[1] Les chercheurs ont baptisé ce constat d'un nom qui reste en mémoire : l'*effet brise-glace* — le chant établit un lien social rapide entre inconnus, sans qu'il y faille de longues conversations personnelles.[2] Leur interprétation évolutionnaire : le chant est vraisemblablement l'un des plus anciens outils de l'humanité pour relier des groupes trop grands pour que chacun puisse se lier d'amitié avec chacun individuellement.[3] Le chœur du mardi soir ne mène aucune expérience sociale. Il se sert d'une machine très ancienne.

6.2 Ce que le chant commun fait aux êtres humains

Le constat du brise-glace n'est pas isolé ; autour de lui s'étend un large champ de recherche sur le chant choral et le bien-être. Des chanteurs d'un chœur amateur, examinés avant et après la répétition, montraient après avoir *chanté* davantage d'humeur positive, moins d'humeur négative — et une hausse de l'immunoglobuline A dans la salive, un marqueur des défenses immunitaires ; la simple *écoute* de musique chorale, dans le même groupe, ne produisait pas cet effet.[4] Si l'on interroge les choristes eux-mêmes, le tableau est unanime : dans l'une des premières enquêtes systématiques, 87 pour cent déclaraient tirer du chant un bénéfice social, 75 pour cent un bénéfice émotionnel.[5] Il faut situer cette littérature avec sobriété — beaucoup

repose sur les déclarations de personnes qui chantent volontairement dans des chœurs, et les études physiologiques sont petites.[6] Mais la convergence est remarquable : à travers les méthodes et les pays, on retrouve le même motif d'humeur rehaussée, de stress abaissé et — surtout — de lien social.

Pour la question de ce livre, un troisième constat est le plus important. La plus grande enquête sur le chant des enfants, l'évaluation du programme national britannique « Sing Up », avec des données portant sur plus de six mille enfants, a trouvé un lien clair : plus le développement vocal d'un enfant est avancé, plus son image de soi et son sentiment d'*appartenir socialement* sont positifs — indépendamment de l'âge, du sexe et de l'origine.[7] Les auteurs appellent eux-mêmes à la prudence : c'est une corrélation, non un sens d'action démontré ; qui se sent appartenir chante peut-être aussi plus volontiers.[8] Mais c'est précisément cette étude qui retient ce dont il s'agit ici : le chant et l'appartenance tiennent l'un à l'autre, et cela indépendamment d'où vient un enfant.

6.3 La voix protégée

On pourrait maintenant objecter : très bien — mais qu'est-ce que cela a à voir avec l'*acquisition d'une langue* ? La réponse ramène au chapitre 2, et elle a deux étages.

Le premier étage est l'anxiété. Aucun moment de l'apprentissage d'une langue n'est aussi chargé d'anxiété que celui où l'on ouvre la bouche devant d'autres : on sonne comme un enfant, on est évalué, et l'image de soi — adulte, compétent, à prendre au sérieux — est en jeu ; la recherche sur l'anxiété langagière a montré que c'est précisément le parler-devant-les-autres qui en est le centre et qu'elle rabat la performance de façon mesurable.[9] Le chœur est la seule forme sociale qui résolve ce problème de façon *constructive* au lieu de l'interdire : la voix singulière est inaudible dans la masse. Celui qui chante une fausse note dans un chœur ne se couvre pas de honte — littéralement, personne ne le remarque. Devient ainsi possible quelque chose qui n'existe nulle part ailleurs : **une production sonore et active dans la langue cible dès le premier jour, avec un risque de honte nul.** Une classe ne peut pas formuler en chœur une lettre administrative ; mais elle peut chanter en chœur, y compris celles et ceux qui ne parlent pas encore un seul mot. Le chœur est, dans la langue du chapitre 2, un abaisseur de filtre construit : il produit la condition sous laquelle l'input arrive et sous laquelle la production ose quelque chose.

La psychologie des organisations a pour cette condition un nom précis. Amy Edmondson l'a appelée la *sécurité psychologique* : la conviction partagée d'un groupe que le risque interpersonnel — demander, se tromper, montrer son ignorance — n'est pas ici sanctionné ; dans son étude portant sur cinquante et une équipes de travail, c'était précisément cette conviction qui était la condition pour que les gens apprennent au lieu de se couvrir.[10] Le chœur

n'a pas à produire cette sécurité par des appels ; il l'intègre dans sa forme : la masse dissimule la voix, donc le risque contre lequel il faudrait être protégé n'existe pas d'emblée.

Le second étage est la persévérance. L'acquisition d'une langue a besoin, c'était la leçon des parties II et III, d'une chose avant tout : la quantité — des mois d'input, des centaines de rencontres avec chaque mot. Mais la quantité n'est pas une grandeur didactique, c'est une grandeur motivationnelle : elle ne naît que là où les gens *reviennent*. Et les gens reviennent là où ils appartiennent. L'effet brise-glace n'est donc pas un bonus bien-être en marge de la méthode, mais son fondement : le lien des premières semaines décide s'il y aura même les mois pendant lesquels les autres mécanismes peuvent agir.

Les deux ensemble forment la chaîne que ce livre affirme — et elle doit être exposée ici aussi ouvertement que possible : la recherche sur les chœurs documente le lien, l'humeur et l'appartenance ; elle ne documente *pas* directement qu'une langue s'acquiert dans un chœur. La conclusion qui va du chœur à l'acquisition passe par deux chaînons intermédiaires documentés — l'anxiété freine l'acquisition, l'appartenance porte la persévérance — et elle est plausible comme chaîne, mais non mesurée comme ensemble.[11] Le chapitre 9 mettra précisément cette chaîne sur la liste de mesure.

6.4 Pour celles et ceux qui doivent arriver

Pour un groupe, cette chaîne n'est pas une figure académique, mais le quotidien : pour les personnes qui sont nouvelles dans un pays. Une réfugiée comme Amal, au chapitre 1, n'a pas seulement un problème de langue ; elle a un problème d'appartenance, et les deux se renforcent mutuellement — qui ne peut pas parler trouve difficilement à se raccorder, et qui n'a aucun raccordement n'a personne avec qui il pourrait parler. Le cours de langue n'interrompt pas ce cercle : elle y est assise exclusivement parmi des personnes qui ont le même problème.

Le chant commun l'interrompt au seul endroit qui soit accessible sans langue. Se tenir dans un chœur à côté de personnes du pays ne présuppose aucun niveau de langue — la mélodie porte, le texte est là, la voix disparaît dans la masse. L'appartenance naît ici *avant* la langue, non comme sa récompense. La recherche sur la community music, qui a accompagné des groupes musicaux composés de personnes réfugiées et de personnes du pays, décrit partout ce même motif : du bien-être, une entente par-dessus les frontières culturelles et un sentiment d'arrivée, bien avant que la langue commune ne porte.[12] C'est le cas rare d'une mesure d'intégration où la société d'accueil n'est pas examinatrice, mais co-chanteuse. La sociologie connaît pour de tels espaces une notion propre : le *tiers-lieu* — ni le domicile, ni le travail ou l'administration, mais le terrain neutre où le statut reste dehors et où la conversation (ici : le chant) est le seul ticket d'entrée.[13] Le chœur est le tiers-lieu qu'une personne qui arrive peut franchir avant de maîtriser la langue des deux premiers.

6.5 Du constat à la construction

Le module chœur de Pontivo traduit ces constats en trois décisions de construction. Premièrement le *mode enseignant/élève* : une personne — cheffe de chœur, enseignant, fondateur du groupe — choisit la chanson, et chez tous les autres apparaît la même, au même moment, avec texte et accords. C'est techniquement insignifiant et socialement décisif : cela produit l'*unique objet commun* que tous regardent en même temps — la forme fondamentale de l'attention partagée dont traite le chapitre suivant. Deuxièmement le *groupe fermé* : chœurs, classes et familles chantent dans des espaces privés que personne ne franchit depuis l'extérieur. Ce qui se passe dans le groupe — la fausse note, la tentative ratée, la chanson qui signifie quelque chose pour quelqu'un — reste dans le groupe ; l'espace protégé qui dissimule la voix singulière dans la sonorité est prolongé au niveau social. Troisièmement le *répertoire* : des chansons simples et riches en répétitions, avec accords (chapitre 5), et parmi elles — cela se révélera dans la partie V comme l'élément le plus dense de tout le concept — des chansons que les membres du groupe apportent eux-mêmes.[14]

Que cette version construite atteigne l'effet des chœurs de salle paroissiale est — règle d'honnêteté — une question ouverte : la recherche citée a étudié des groupes physiquement rassemblés, non des groupes assistés par une application.[15] Mais les décisions de construction ne sont pas des inventions, ce sont des traductions : chacune d'elles a derrière elle un constat que l'on peut nommer.

Reste la question de savoir qui conduit le groupe — et ce que devient la langue maternelle lorsqu'un enseignant pilote une classe à travers le film et la chanson. C'est la salle de classe, et elle a sa propre théorie, vieille de cent ans.

Notes

1. Pearce, E./Launay, J./Dunbar, R. I. M., The ice-breaker effect: singing mediates fast social bonding, *Royal Society Open Science* 2 (2015) 150221, doi:10.1098/rsos.150221 (libellé du résumé vérifié sur la notice PubMed) : « we followed newly formed singing and non-singing (crafts or creative writing) adult education classes over seven months » ; moments de mesure : mois 1, 3 et 7 ; « Singers demonstrated a significantly greater increase in closeness at timepoint 1 ». ↩
2. Ibid. : « Although singers and non-singers felt equally connected by timepoint 3, singers experienced much faster bonding » — la notion d'« ice-breaker effect » vient du titre de l'étude. Par souci d'honnêteté, il faut ajouter ceci : après sept mois, les groupes étaient à égalité ; l'avantage du chœur est un avantage de *vitesse*, non d'état final. Pour des contextes d'apprentissage où les premières semaines décident de l'abandon ou de la persévérance, c'est précisément la vitesse qui est la grandeur pertinente (à ce sujet, 6.3). ↩

3. Sur le cadrage évolutionnaire (le chant comme mécanisme de lien pour de grands groupes, « social bonding hypothesis »), la discussion chez Pearce et al. (note 1) ainsi que les travaux de Dunbar sur les tailles de groupe et le toilettage social (parmi bien d'autres : Dunbar, *Grooming, Gossip, and the Evolution of Language*, Cambridge/MA 1996, à vérifier sur l'imprimé). L'interprétation évolutionnaire est une théorie d'arrière-plan, non une preuve portante de ce chapitre. ↩
4. Kreutz, G./Bongard, S./Rohrmann, S./Hodapp, V./Grebe, D., Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state, *Journal of Behavioral Medicine* 27 (2004) 623–635 (résumé vérifié sur la notice PubMed) : membres d'un chœur amateur, deux répétitions à une semaine d'intervalle, condition chanter vs écouter de la musique chorale ; « Singing leads to increases in positive affect and S-IgA, while negative affect is reduced. » ↩
5. Clift, S./Hancox, G., The perceived benefits of singing: findings from preliminary surveys of a university college choral society, *Journal of the Royal Society for the Promotion of Health* 121 (2001) 248–256 (résumé vérifié sur la notice PubMed) : n = 84 et 91 ; 87 % d'adhésion au bénéfice social, 75 % au bénéfice émotionnel ; six dimensions de bénéfice. ↩
6. Sur la sobriété qui s'impose : l'autosélection (qui aime les chœurs chante dans des chœurs), les formats déclaratifs et de petits échantillons physiologiques limitent la force probante de la littérature sur le bien-être ; les constats sur le lien (note 1) sont méthodologiquement plus solides (longitudinal, groupes de comparaison). Ce livre appuie donc sa chaîne (6.3) prioritairement sur le lien et l'anxiété, non sur des effets de santé. ↩
7. Welch, G. F./Himonides, E./Saunders, J./Papageorgi, I./Sarazin, M., Singing and social inclusion, *Frontiers in Psychology* 5 (2014) 803, doi:10.3389/fpsyg.2014.00803 (accès libre, références vérifiées sur le texte intégral PMC) : 6 087 enfants avec données appariées (sur 11 258 étudiés) : « The higher the normalized singing development rating, the more positive the child's self-concept and sense of being socially included, irrespective of singer age, sex and ethnicity. » ↩
8. Ibid. : « Whilst caution is needed in the interpretation of the correlation and regression data, and in the pitfalls of ascription of causal effects between measures … » — les auteurs eux-mêmes laissent le sens d'action ouvert ; ce livre reprend donc le constat seulement comme un lien, non comme une preuve. ↩
9. Horwitz/Horwitz/Cope 1986 (chapitre 2, note 9), p. 125 et s. ; sur la place à part du parler, ibid. Les items de la FLCAS tournent essentiellement autour de situations de parole devant la classe. ↩
10. Edmondson, A., Psychological Safety and Learning Behavior in Work Teams, *Administrative Science Quarterly* 44 (1999) 350–383 : la sécurité psychologique comme « a shared belief held by members of a team that the team is safe for interpersonal risk taking » (définition et dispositif d'étude — 51 équipes de travail d'une entreprise manufacturière — vérifiés sur un miroir du texte intégral librement accessible ; à vérifier

sur l'imprimé). Avertissement sur la transposition : les constats proviennent d'équipes de travail, non de groupes d'apprentissage ; ce qui est repris, c'est le construit et son mécanisme (sécurité → comportement d'apprentissage), non une taille d'effet pour des chœurs. ↩

11. Avertissement sur la transposition, en toute forme : sont documentés (a) le lien rapide par le chant de groupe (note 1 s.), (b) le rapport entre chant et appartenance (note 7), (c) l'effet dépressif de l'anxiété de parole sur la performance (note 9, chapitre 2 note 10) et (d) la nécessité de grandes quantités d'input (partie II). N'est PAS directement documentée la conclusion d'ensemble « module chœur → meilleure acquisition de la langue » ; c'est une inférence à deux étages passant par le filtre et la persévérance, et elle est formulée au chapitre 9 comme tâche de mesure (mesures d'anxiété, d'appartenance et de rétention en mode chœur). ↩
12. Parmi bien d'autres : Dow, R./Vougioukalou, S., Wellbeing and Integration through Community Music: The Role of Improvisation in a Music Group of Refugees, Asylum Seekers and Local Community Members, *Contemporary Music Review* 38 (2019), doi:10.1080/07494467.2019.1684075 (à vérifier sur le texte intégral ; accès éditeur bloqué). Cette littérature est majoritairement qualitative (entretiens, observation) et est présentée ici comme un motif convergent, non comme une preuve d'effet. ↩
13. Oldenburg, R., *The Great Good Place*, New York 1989 (à vérifier sur l'imprimé) : le « tiers-lieu » comme espace communautaire neutre, niveleur de statut, à côté du domicile (premier lieu) et du travail (deuxième lieu) — élaboré à partir des cafés, des bistrots, des salons de coiffure. Avertissement sur la transposition : les critères d'Oldenburg (neutralité, faible seuil d'accès, la conversation comme activité principale, un public d'habitués) s'appliquent au chœur dans une mesure remarquable — avec cette différence qu'à la place de la conversation vient le chant ; employé comme situation sociologique, non comme preuve. ↩
14. Note fonctionnelle : le module chœur est en service (pontivo.de/choere, état au 15 octobre 2026) : groupes fermés, mode enseignant/élève, suivi du texte en tant qu'invité par code d'invitation, sans compte. ↩
15. En outre, question ouverte et nommée par souci de probité : savoir si un chant commun médié de façon synchrone par une plateforme (même texte, même moment, le cas échéant dans des espaces distribués) porte les effets de lien des chœurs physiques n'est pas exploré ; la physique de la latence interdit de toute façon le chant distribué en temps réel — le module de Pontivo est donc construit prioritairement pour l'*espace commun* (classe, répétition de chœur, famille), la plateforme livre le texte, les accords et la simultanéité, non l'espace. ↩

Chapitre 7 · La salle de classe : attention partagée, langue maternelle légitimée

7.1 L'interdiction

Dans le cours d'intégration d'Amal, il y avait une règle qui paraissait si évidente que personne ne la justifiait : *Ici, on ne parle qu'allemand.* L'allemand était expliqué en allemand, à des personnes qui ne savaient pas encore l'allemand ; l'arabe qu'Amal apportait avec elle — une langue pleinement développée, dans laquelle elle peut penser, argumenter, être ironique — n'était pas un outil dans la salle de cours, mais un incident. Cette interdiction a une histoire, et elle est plus courte et plus prosaïque que son prestige ne le laisse supposer : la « méthode directe » des écoles de langues du début du vingtième siècle a fait du monolinguisme un principe — notamment parce qu'il était pratique pour des chaînes d'écoles au personnel enseignant international : qui n'enseigne que dans la langue cible n'a pas besoin de savoir les langues de ses élèves.[1] De la nécessité est née une doctrine, de la doctrine un réflexe. Scientifiquement, il n'a jamais été particulièrement bien fondé.

Ce chapitre traite de ce qu'est en réalité une salle de classe lorsqu'on la pense depuis la recherche sur l'apprentissage — et de la raison pour laquelle la langue maternelle n'y est pas l'ennemie de la langue cible, mais son échafaudage le plus important.

7.2 L'échafaudage

Le fondement théorique vient des années 1930, de Lev Vygotski, et il a un noyau simple : apprendre est d'abord un processus social et ne devient qu'ensuite un processus intérieur. Ce qu'un enfant ne réussit aujourd'hui qu'en tenant la main de quelqu'un qui sait, il le réussira demain seul ; l'intervalle entre les deux — entre ce que l'on sait faire de façon autonome et ce que l'on sait faire « sous conduite ou en coopération avec des partenaires plus capables » — Vygotski l'a appelé la *zone proximale de développement*.[2] L'apprentissage a lieu dans cette zone, et la tâche de l'enseignant n'est pas de tendre du savoir par-dessus, mais d'exploiter la zone : poser les tâches de telle sorte qu'elles se situent un pas au-dessus de ce qui est déjà su seul, et donner exactement assez d'appui pour que le pas réussisse. Pour cet appui, Wood, Bruner et Ross ont forgé l'image qui circule depuis dans toutes les didactiques : l'*échafaudage* (scaffolding) — la structure de soutien que le plus expérimenté érige et démonte pièce par pièce, tandis que l'apprenant porte ce qu'il peut déjà porter.[3]

On remarquera avec quelle exactitude cela recouvre le *i + 1* de Krashen, au chapitre 2, qui vient d'une tradition toute différente : les deux disent que l'apprentissage se produit au bord du

pouvoir, un pas au-dessus du niveau — et que c'est une *aide à la compréhension*, non une explication de la forme, qui porte par-dessus ce bord. Le sous-titre bilingue du chapitre 3 n'est rien d'autre qu'un échafaudage au sens propre : un appui qui maintient la compréhension là où la force langagière propre ne suffit pas encore, et que l'on peut démonter — ligne par ligne — lorsqu'il n'est plus nécessaire.

7.3 La langue maternelle comme échafaudage

Mais avec quoi construit-on des échafaudages ? L'appui le plus stable que possède un apprenant adulte est la langue qu'il sait déjà. Wolfgang Butzkamm — l'angliciste allemand qui a écrit toute une vie de chercheur contre le dogme du monolinguisme — l'a ramené à la forme la plus simple : la langue maternelle n'est pas le facteur de perturbation de l'apprentissage des langues étrangères, mais son plus grand capital ; c'est en elle que nous avons appris ce qu'est parler en général — que des sons signifient, que des phrases se construisent, que l'on agit avec des mots — et ce savoir migre dans chaque langue suivante.[4] La recherche plus récente sur le plurilinguisme a déployé la même idée sous la notion de *translanguaging* : les personnes plurilingues ne possèdent pas deux casiers linguistiques séparés, mais un répertoire commun dans lequel elles puisent selon la situation ; un enseignement qui met la première langue à la porte met à la porte une partie de l'outil de pensée de ses élèves.[5]

Ce n'est pas une simple position théorique ; les traces empiriques sont dispersées à travers tout ce livre, et il vaut la peine de les rassembler une fois. La méta-analyse sur les gloses, au chapitre 4 : les explications de mots dans la *langue maternelle* battent les explications dans la langue cible, y compris chez les apprenants avancés.[6] L'étude oculométrique du chapitre 3 : la ligne de sous-titre en langue maternelle est l'ancre stable depuis laquelle les apprenants explorent la ligne en langue cible — le sens repose, la forme est conquise.[7] Et le déluge de livres du chapitre 2 a agi dans des écoles dont les enfants parlaient le fidjien ou le hindi à la maison — personne ne leur a extirpé cela, l'anglais est venu *en plus*.[8] Partout où l'on admet la langue maternelle comme un outil au lieu de la traiter comme une fuite, elle travaille avec nous.

Pour qu'aucun malentendu ne naisse : rien de tout cela ne parle contre le fait d'avoir le plus de langue cible possible dans la pièce — la quantité d'input compréhensible reste la première de toutes les conditions. Cela parle contre l'*interdiction* de la première langue : contre le cours qui refuse une aide à la compréhension qui coûterait une demi-phrase, et contre l'humiliation de celles et ceux qui pensent dans leur propre langue. L'échafaudage ne se met pas en travers du bâtiment. Il le rend possible.

Celui qui tiendrait cela pour une position marginale sous-estime à quel point la politique linguistique en est déjà loin : le Conseil de l'Europe a expressément élevé le plurilinguisme au rang de ligne directrice pédagogique dans le volume complémentaire du Cadre de référence

(2020) — la compétence plurilingue n'y vaut pas comme la somme de monolinguismes séparés, mais comme un répertoire cohérent que l'enseignement doit utiliser au lieu de le réprimer.[9] Intégrer la première langue n'est donc plus une audace didactique, mais la direction que désigne le cadre européen lui-même.

7.4 La salle de classe construite

La salle de classe de Pontivo met en œuvre techniquement ces deux idées — attention partagée sous conduite, langue maternelle comme échafaudage — et, sur un point, elle fait quelque chose qu'une salle de classe physique ne peut pas faire.

Le *mode enseignant/élève* est la forme construite de l'attention partagée : l'enseignant choisit — le film, la scène, la chanson — et chez tous apparaît la même chose au même moment. Cela paraît banal et c'est le noyau de toute forme d'enseignement depuis l'école de village : un objet, de nombreux regards, une main qui conduit. La plateforme n'enlève pas à l'enseignant la conduite, mais la friction — la distribution, le « page 34, deuxième paragraphe », le « chez moi, ça ne marche pas ».[10]

La particularité tient au second niveau : le même film tourne pour chaque élève *avec sa propre ligne en langue maternelle*. Une classe où l'arabe, l'ukrainien, le tigrigna et le polonais sont assis ensemble — le cas normal de toute classe d'intégration — regarde le même film allemand, et chaque enfant a son propre échafaudage en dessous. Dans la salle de classe physique, c'est impossible ; aucun enseignant ne peut chuchoter en quatre langues à la fois. Ici, c'est une ligne de configuration : le moteur sépare la langue cible, la langue maternelle et l'interface comme des axes indépendants.[11] C'est du translanguaging comme infrastructure : la classe partage l'objet et la langue cible — et chacun garde son propre outil.

Et l'enseignant ? Il n'est pas remplacé, il est *déchargé* — précisément de la part de son travail que la machine sait faire (assurer la compréhension en quatre langues simultanément), au profit de la part que lui seul sait faire : choisir, interpréter, encourager, conduire la conversation après le film, poser la tâche qui fait de ce qui a été vu de la langue. La recherche en éducation donne raison à ce partage du travail : dans la grande synthèse de recherche de Hattie, la qualité de la relation entre l'enseignant et les apprenants figure parmi les facteurs d'influence les plus forts que l'on connaisse sur la réussite de l'apprentissage — du temps qu'une machine rend au travail relationnel est donc investi dans la grandeur la plus efficace de l'enseignement.[12] Un outil censé rendre les enseignants superflus serait aussi mal construit du point de vue pédagogique : le principe de l'échafaudage *a besoin* de celui qui sait, qui sait quand il soutient et quand il lâche.[13]

À cet endroit, la probité exige aussi une déclaration d'intention, signalée comme telle : à terme, la pratique doit codévelopper la plateforme — des enseignants qui signalent une explication au

clic imprécise ou qui versent au fonds commun une traduction de chanson réussie, sur le modèle de la recherche-action participative.[14] Cette boucle de rétroaction n'est pas encore construite ; elle figure ici comme plan de construction, non comme affirmation.

7.5 Ce qui manque encore à la salle de classe

Une salle de classe qui regarde ensemble et chante ensemble a beaucoup — mais elle n'a encore aucune matière qui vienne *de la classe elle-même*. Jusqu'ici, tout le matériau coule dans une seule direction : du catalogue vers le groupe. L'élément le plus dense de tout le concept retourne cette direction : il fait de ce que les apprenants apportent — les chansons de leur origine — le matériau du cours, et des apprenants eux-mêmes ceux qui l'expliquent. Ce qui se passe alors, sur le plan langagier, social et humain, est le chapitre suivant — et avec lui, ce livre en arrive à sa pointe.

Notes

1. Sur l'histoire de la méthode directe et du principe de monolinguisme (écoles Berlitz, mouvement de réforme) : Richards, J. C./Rodgers, T. S., *Approaches and Methods in Language Teaching*, 3e éd., Cambridge 2014, chap. sur la Direct Method (à vérifier sur l'imprimé) ; de façon critique sur la base scientifique du dogme, Butzkamm/Caldwell (note 4), chap. 1. L'histoire causale abrégée donnée dans le texte (pratique de recrutement → principe) suit l'exposé de Butzkamm et reste à vérifier comme telle. ↩
2. Vygotski, L. S., *Mind in Society: The Development of Higher Psychological Processes*, Cambridge/MA 1978, p. 86 : la zone proximale de développement comme « the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers » (notre traduction dans le texte ; page selon l'édition standard, à vérifier sur l'imprimé). ↩
3. Wood, D./Bruner, J. S./Ross, G., The role of tutoring in problem solving, *Journal of Child Psychology and Psychiatry* 17 (1976) 89–100 (à vérifier sur l'imprimé ; accès éditeur bloqué). La notion de « scaffolding » pour la fonction de soutien démontable du tuteur vient de cet article ; la restitution suit la réception standard. ↩
4. Butzkamm, W./Caldwell, J. A. W., *The Bilingual Reform: A Paradigm Shift in Foreign Language Teaching*, Tübingen 2009 (à vérifier sur l'imprimé) ; thèse centrale : « The mother tongue is the greatest asset people bring to the task of foreign language learning » (formulation standard des auteurs, référence à vérifier sur l'imprimé). En allemand : Butzkamm, *Psycholinguistik des Fremdsprachenunterrichts*, 3e éd., Tübingen 2002 (à vérifier sur l'imprimé). ↩
5. García, O., *Bilingual Education in the 21st Century: A Global Perspective*, Malden 2009 (à vérifier sur l'imprimé) ; García, O./Wei, L., *Translanguaging: Language,*

Bilingualism and Education, Basingstoke 2014 (à vérifier sur l'imprimé). Avertissement sur la transposition : le translanguaging est bien élaboré comme description de la pratique plurilingue et comme posture pédagogique, mais ce n'est pas un principe d'action quantifié expérimentalement ; les appuis quantitatifs de ce chapitre sont les constats particuliers rassemblés aux notes 6–8. ↩

6. Yanagisawa/Webb/Uchihara 2020 (chapitre 4, note 5 s.) : « L1 glosses yielded greater learning than L2 glosses », sans interaction avec le niveau de langue. ↩
7. Liao/Kruger/Doherty 2020 (chapitre 3, note 13) : attention du regard plus stable sur la ligne en L1, exploration flexible de la ligne en L2, aucune surcharge due à la double piste. ↩
8. Elley/Mangubhai (chapitre 2, note 5 s.) ; sur le débordement des effets vers la première langue (le fidjien) la deuxième année, l'exposé des auteurs sur place (Fiji Intermediate Examination, moyenne de 58 % également dans la matière Fijian Language). ↩
9. Council of Europe, *Common European Framework of Reference for Languages: Learning, teaching, assessment — Companion volume*, Strasbourg 2020 : la compétence plurilingue et pluriculturelle comme concept clé et ligne directrice de l'éducation aux langues (énoncé central vérifié sur les pages officielles du Conseil de l'Europe ; références précises à vérifier sur le texte intégral PDF — l'accès direct à rm.coe.int était bloqué électroniquement). Sciemment désigné comme « ligne directrice » : le cadre de référence est un cadre, non une norme juridique. ↩
10. Note fonctionnelle : la salle de classe est en service (pontivo.de/fuer-lehrer, état au 15 octobre 2026) : classe avec code d'invitation, devoirs, chat de classe, cinéma synchronisé commun ; affichage sur vidéoprojecteur ou téléviseur via pontivo.de/tv. ↩
11. L'architecture en axes (langue cible / langue maternelle / langue de l'interface, indépendantes pour chaque utilisateur) est une description de produit, non un constat de recherche ; elle est rapportée ici parce qu'elle est la condition technique de la classe plurilingue. Savoir si la classe plurilingue avec ligne d'échafaudage individuelle a l'avantage attendu par rapport à une ligne unique est non mesuré (chapitre 9). ↩
12. Hattie, J., *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*, Londres 2009 (à vérifier sur l'imprimé) : la relation enseignant-apprenants, avec d ≈ 0,72, parmi les facteurs d'influence les plus forts — expressément pas « le plus fort » : dans les classements ultérieurs de Hattie, d'autres facteurs la dépassent (par exemple l'efficacité collective des enseignants, d ≈ 1,57 ; valeurs d'après les panoramas tenus par l'auteur, à vérifier sur l'imprimé). De plus, la méthodologie de Hattie (agrégation de méta-analyses hétérogènes) a été substantiellement critiquée — l'ordre de grandeur « la relation agit fortement » passe néanmoins pour robuste ; rien de plus n'est affirmé ici. ↩
13. Sur l'orientation par les tâches (la tâche après l'input, celle qui fait de la compréhension de la langue) : Ellis, R., *Task-based Language Learning and Teaching*, Oxford 2003 ; Long, M., *Second Language Acquisition and Task-Based Language Teaching*, Malden 2015 (tous deux à vérifier sur l'imprimé). Pontivo livre sciemment la part input et la part

communauté et laisse le niveau des tâches à l'enseignant — un partage du travail, non une prétention à l'exhaustivité. ↩

14. Sur le modèle méthodologique : Kemmis, S./McTaggart, R./Nixon, R., *The Action Research Planner: Doing Critical Participatory Action Research*, Singapour 2014 (à vérifier sur l'imprimé). Le statut est signalé dans le texte : intention de construction déclarée, non fonctionnalité existante — la boucle de rétroaction (correction des gloses, versement de strophes) n'existe pas actuellement dans la plateforme. ↩

Partie V — L'épreuve

> „Ihre Gesänge sind das Archiv des Volks, der Schaz ihrer Wissenschaft und Religion … Abdruck ihres Herzens, Bild ihres häuslichen Lebens in Freude und Leid, beym Brautbett und Grabe."
> — Johann Gottfried Herder, Deutsches Museum 1777, S. 433

Chapitre 8 · Coro One Sky : la chanson venue du pays d'origine, chantée dans la nouvelle langue

8.1 L'exercice

L'exercice dont traite ce chapitre se décrit en une phrase : un membre du chœur apporte une chanson de son pays d'origine ; le groupe la traduit ensemble dans la nouvelle langue ; puis tous la chantent — en italien, en allemand — sur l'ancienne mélodie.

Il faut se représenter le déroulement concrètement pour voir ce qu'il contient. Un adolescent syrien fredonne la mélodie que sa mère a chantée ; il écrit les vers arabes et explique au groupe, vers par vers, ce qu'ils signifient — ce qu'ils signifient *vraiment*, car le sens du dictionnaire lui paraît trop peu. Le groupe propose des versions italiennes, rejette, affine : ce vers est trop long pour la mélodie, ce mot ne tombe pas sur la note, ici l'accent doit aller sur la deuxième syllabe. L'enseignant aide avec la grammaire là où cela coince. Après deux soirées, une strophe italienne chantable est là. Alors le chœur se lève et la chante — et le garçon dont c'est la chanson entend vingt personnes chanter dans la nouvelle langue ce que sa mère chantait dans l'ancienne.

Dans cet exercice discret, tous les fils de ce livre convergent en un seul point. Il vaut la peine de démonter la mécanique — ce sont cinq mécanismes, et pour chacun d'eux la preuve est déjà prête dans un chapitre antérieur.

8.2 Cinq mécanismes

Premièrement : la mélodie est déjà là. Le chapitre 5 s'achevait sur une liste de conditions : la chanson porte le texte lorsque la mélodie est simple, qu'elle se répète et qu'elle n'a pas elle-même à être apprise — le contre-constat de Racette est précisément né de ce que ses sujets devaient maîtriser à la fois le texte *et* une mélodie inconnue.[1] La chanson du pays d'origine est le cas limite où cette condition est parfaitement remplie : la mélodie est en place depuis l'enfance, avec son rythme, ses longueurs de vers et ses schémas d'accentuation — tout l'échafaudage qui, selon Wallace, « segmente » le texte est déjà prêt dans la mémoire.[2] Les nouveaux mots italiens sont accrochés à une structure ancienne, profondément ancrée et chargée d'émotion. La recherche sur la mémoire ne connaît pas d'emplacement de stockage plus favorable pour une langue étrangère.

Deuxièmement : traduire est la forme la plus profonde du travail sur les mots. Le chapitre 4 a nommé les trois ingrédients qui font qu'un mot adhère : le besoin, la recherche, l'évaluation.[3] La traduction commune d'une chanson pousse les trois à leur maximum. Le besoin n'est pas simulé didactiquement, il est réel — c'est *ma* chanson, et la version doit être juste ; la recherche passe par le dictionnaire, l'enseignant et la discussion de groupe ; et l'évaluation est la plus sévère que l'on puisse concevoir : le mot ne doit pas seulement être correct, il doit être *chantable* — nombre de syllabes, accentuation, sonorité. Chaque vers est une glose à choix multiple sous forme collective : produire des versions, les comparer, les rejeter.[4] Et ce qui se tient là à la fin, le groupe l'a *produit lui-même* — l'effet de génération, l'un des constats les plus robustes de la psychologie de la mémoire, dit exactement ceci : ce que l'on a produit soi-même est mieux retenu que ce qui a seulement été lu, et cela à travers toutes les formes de test.[5]

Troisièmement : les rôles s'inversent. Dans le cours de langue, la personne réfugiée est structurellement celle qui est en déficit — celle à qui il manque quelque chose. Dans cet exercice, elle est, pour la durée de la chanson, le seul expert de la pièce : elle connaît l'original, elle décide si la traduction est juste, l'enseignant et le groupe dépendent d'elle. L'anxiété langagière, dont le chapitre 2 a documenté l'effet dépressif sur la performance, tient à l'image de soi menacée — à l'adulte qui s'entend sonner comme un enfant.[6] Ici, c'est le contraire qui se produit : la compétence devient visible avant que la nouvelle langue puisse l'exprimer. L'étude Sing Up du chapitre 6 en a donné une trace empirique : un chant qui réussit est lié à l'image de soi et au sentiment d'appartenance — les auteurs supposent comme mécanisme précisément l'efficacité personnelle éprouvée.[7]

Quatrièmement : l'intégration va dans les deux sens. L'assimilation dit : laisse tes chansons derrière toi et apprends les nôtres. Cet exercice dit le contraire : *Ta chanson mérite d'exister dans la nouvelle langue.* L'origine n'est pas déposée, elle est portée de l'autre côté ; la nouvelle langue devient le support de l'histoire propre au lieu d'en être la concurrente. C'est du translanguaging non comme technique d'enseignement, mais au sens propre : deux langues qui travaillent à la même strophe.[8] Et la société d'accueil n'est pas examinatrice, mais co-traductrice et co-chanteuse — elle apprend la chanson, l'histoire derrière elle, un morceau des personnes avec qui elle chante. De tous les formats que décrit ce livre, c'est le seul dans lequel les personnes du pays *apprennent* quelque chose de celles qui arrivent — l'intégration cesse d'être une voie à sens unique.

Que précisément cette constellation défasse les préjugés n'est pas ici un pieux espoir, mais l'un des constats les mieux éprouvés de la psychologie sociale. L'hypothèse du contact de Gordon Allport nomme les conditions sous lesquelles une rencontre entre groupes agit : un statut égal dans la situation et un but commun — et la grande méta-analyse de Pettigrew et Tropp, portant sur plus de cinq cents études, le confirme : le contact abaisse typiquement les préjugés de toute façon, mais c'est sous les conditions d'Allport qu'il le fait le plus fortement.[9] Que l'on lise ces

conditions à côté de l'exercice : l'inversion des rôles établit le statut égal, la strophe chantable est le but commun. Ce que la musique apporte en plus — un rythme partagé, un souffle partagé, et dans la chanson traduite l'expérience audible que la chanson venue d'Asmara raconte les mêmes espoirs que celle du pays — est discuté en cognition musicale comme un mécanisme de lien social ; on ne doit pas le présenter comme un effaceur de préjugés neurobiologiquement assuré, mais bien comme un amplificateur plausible des conditions du contact.[10]

Cinquièmement : l'exécution est sans anxiété. À la fin, on chante — la seule forme de production sonore dans la nouvelle langue qui soit possible dès le premier jour sans risque de honte, avec la voix singulière protégée dans la masse (chapitre 6). Et la chanson traduite ne disparaît pas : elle entre dans le répertoire du groupe et est chantée encore et encore — la répétition dont l'apprentissage incident a besoin (chapitre 4) est intégrée, sans jamais s'appeler « exercice ».

On le voit : aucun de ces cinq mécanismes n'est une invention de ce livre — échafaudage mélodique, profondeur de traitement, effet de génération, efficacité personnelle, abaissement de l'anxiété sont, chacun pour soi, des constats documentés, pour certains vieux de plusieurs décennies. L'exercice est leur réunion dans un seul format. C'est précisément là sa force — et, à vrai dire, aussi son problème de preuve, sur lequel revient la dernière section.

8.3 Le récit : pourquoi l'ensemble a besoin d'une histoire

Pontivo raconte cet exercice — et la méthode derrière lui — comme une histoire : *Coro One Sky*, une bande dessinée sur un chœur dans lequel des enfants de différents pays chantent ensemble sous le même ciel. On pourrait y voir un emballage. La recherche en cognition suggère que c'est davantage.[11]

Jerome Bruner — celui-là même de qui vient l'image de l'échafaudage, au chapitre 7 — a soutenu que la pensée narrative est un mode de compréhension propre et irremplaçable : les êtres humains ordonnent la réalité sociale en histoires, non en phrases dotées d'une valeur de vérité.[12] Mar et Oatley en ont fait un programme de recherche : la fiction fonctionne comme une *simulation* du monde social — elle abstrait, condense et fait vivre au lecteur une expérience sociale qu'il n'a pas faite lui-même.[13] Pour la didactique des langues s'y ajoute un constat plus tangible, déjà apparu au chapitre 2 : le déluge de livres a agi avec des livres d'*histoires*, et la tradition de la lecture extensive a montré depuis, encore et encore, qu'une lecture narrative, librement choisie et facile — y compris expressément les bandes dessinées — porte la langue, parce que l'histoire tire le lecteur là où l'exercice devrait le pousser.[14]

Dans le concept, le récit prend donc en charge trois tâches. Il donne à la méthode des *figures et un ordre* : qui lit la bande dessinée voit des enfants qui vivent exactement ce que l'exercice

exige — l'étrangeté, le premier chant en commun, la chanson qui est la sienne — et sait par là ce qui vient ; la prévisibilité est une aide à la compréhension, pour les enfants comme pour l'enseignante qui ose ce format pour la première fois. Il livre une *matière de lecture facile* dans la langue cible, doublement codée au sens propre — image et texte —, comme la forme d'input au seuil le plus bas de la plateforme. Et il rend la méthode *racontable* : une école, une réunion de parents, un chef de chœur comprend « un chœur sous un même ciel » plus vite que n'importe quelle liste d'hypothèses.

Ici encore, la limite d'honnêteté : les données narratives proviennent de la recherche sur la lecture et la fiction ; que la *bande dessinée d'une plateforme* remplisse ces fonctions est une transposition, non une preuve.[15] La bande dessinée est la pièce la moins mesurable du concept — et en même temps la moins chère : elle ne coûte aucun temps de cours et ne concurrence rien. Des paris à faible enjeu et à gain plausible, on peut les prendre, à condition de les présenter comme des paris.

8.4 La règle de prudence — et l'unique grand pari

Deux choses doivent se tenir à la fin de ce chapitre, précisément parce qu'il est le plus enthousiaste du livre.

La première est une règle de protection. Les chansons du pays d'origine ne sont pas neutres : à elles tiennent l'enfance, la famille, parfois la perte et pire encore. Un format qui invite des personnes à faire de leur part la plus intime un matériau de cours a besoin du volontariat comme principe dur — celui à qui la chanson appartient décide *seul* si on la chante, et celui qui ne veut pas (encore) chante avec les autres leurs chansons, et ce n'est pas un déficit. Et personne n'est nommé représentant de « sa culture » : le garçon apporte *sa* chanson, non « la musique syrienne ». La frontière entre la reconnaissance et la mise en spectacle est mince, et elle passe là où se situe la décision : chez le chanteur, non chez le format.[16]

La seconde est le bilan de l'état des preuves, sans argument de défense : les cinq mécanismes sont documentés isolément ; leur combinaison dans ce format ne l'est pas. À la connaissance de ce livre, la chanson du pays d'origine traduite n'a pas été étudiée empiriquement comme format intégré d'apprentissage et d'intégration — c'est le pari de conception central de tout le concept, celui dont le fondement de justification est le meilleur et les preuves directes les plus minces.[17] Un concept qui se dit scientifique n'a pour cela qu'un seul traitement probe : nommer le pari, garder l'enjeu petit, planifier la mesure. Ce qu'il faudrait mesurer exactement — et ce que le concept dans son ensemble peut et ne peut pas —, c'est le dernier chapitre.

Notes

1. Racette/Peretz 2007 (chapitre 5, note 7) : ce qui y était appris, c'étaient des chansons non familières — la mélodie faisait partie de la matière à apprendre, le chant était une double tâche. Le format de la chanson du pays d'origine élimine précisément cette double tâche pour celui qui apporte la chanson ; pour les autres membres du groupe, la mélodie est apprise après quelques passages (à supposer des chansons simples et strophiques, chapitre 5, notes 8 et 11). ↩
2. Wallace 1994 (chapitre 5, note 8) : la mélodie soutient le rappel du texte lorsqu'elle revient et qu'elle est facile à apprendre ; sa performance d'articulation (longueurs de vers, schémas d'accentuation, découpage en blocs) se transmet à tout texte qui vient s'y ajuster — donc aussi à la traduction. Que le transfert de l'appui mnésique de la version originale à la version traduite du texte réussisse sans perte est une attente plausible, non un constat éprouvé. ↩
3. Laufer/Hulstijn 2001 (chapitre 4, note 8) : need, search, evaluation. ↩
4. Sur la supériorité des formes de glose actives en traitement, Yanagisawa/Webb/Uchihara 2020 (chapitre 4, notes 5 et 7) ; l'analogie vers de chanson ↔ glose à choix multiple est une interprétation de ce livre. ↩
5. Slamecka, N. J./Graf, P., The generation effect: Delineation of a phenomenon, *Journal of Experimental Psychology: Human Learning and Memory* 4 (1978) 592–604 (énoncé central du résumé vérifié : supériorité de la condition « generate » sur la condition « read » dans les cinq expériences, à travers le rappel libre et assisté, la reconnaissance et les jugements de confiance ; à vérifier sur le texte intégral). Avertissement sur la transposition : l'effet de génération classique porte sur des mots isolés ou des paires de mots en laboratoire ; son extension à des vers de chanson produits collectivement est une analogie. ↩
6. Horwitz/Horwitz/Cope 1986 (chapitre 2, note 9) : l'anxiété langagière comme menace pour l'image de soi et l'expression de soi ; chapitre 2, note 10 sur l'effet sur la performance. ↩
7. Welch et al. 2014 (chapitre 6, note 7 s.) : lien entre développement vocal, image de soi et appartenance ; comme mécanisme, les auteurs supposent « perceived competence and singing self-efficacy » — expressément comme une inférence, non comme une preuve. ↩
8. García/Wei 2014 (chapitre 7, note 5). La traduction de la chanson est en même temps le point où la supériorité, citée au chapitre 7, des gloses en L1 (chapitre 4, note 6) devient pratique : la clarification du sens passe constitutivement par la première langue de celui qui apporte la chanson. ↩
9. Allport, G. W., *The Nature of Prejudice*, Reading/MA 1954, en particulier le chap. 16 (à vérifier sur l'imprimé) ; de façon méta-analytique, Pettigrew, T. F./Tropp, L. R., A meta-analytic test of intergroup contact theory, *Journal of Personality and Social Psychology* 90 (2006) 751–783 (constats centraux vérifiés sur le résumé) : 713 échantillons indépendants issus de 515 études ; le contact réduit typiquement les préjugés même sans les conditions optimales d'Allport, et plus fortement avec elles — les conditions sont favorisantes, non impérativement nécessaires. C'est précisément pourquoi il n'est pas

affirmé ici que l'exercice serait la seule forme efficace — seulement qu'il établit les conditions favorisantes de façon inhabituellement complète. ↩

10. Sur la discussion de la musique comme mécanisme de lien social, parmi bien d'autres Koelsch, S., *Brain and Music*, Chichester 2012 (à vérifier sur l'imprimé), ainsi que les constats sur le lien au chapitre 6 (notes 1 s., 4). Est expressément écartée la lecture plus forte selon laquelle faire de la musique ensemble dissoudrait « les frontières neurobiologiques entre l'in-group et l'out-group » — l'état des constats ne couvre pas cela. Pour qui préfère saisir ce moment de façon sociologique : la notion de *résonance* de Hartmut Rosa (Rosa, *Resonanz. Eine Soziologie der Weltbeziehung*, Berlin 2016, à vérifier sur l'imprimé) le décrit comme un rapport au monde réussi — cité comme cadre d'interprétation, non comme preuve. ↩
11. Note fonctionnelle : la bande dessinée est lisible sur la plateforme (pontivo.de/comic, état au 15 octobre 2026). ↩
12. Bruner, J., *Actual Minds, Possible Worlds*, Cambridge/MA 1986, chap. 2 (mode paradigmatique vs mode narratif ; à vérifier sur l'imprimé) ; id., The narrative construction of reality, *Critical Inquiry* 18 (1991) 1–21 (à vérifier sur le texte intégral). ↩
13. Mar, R. A./Oatley, K., The function of fiction is the abstraction and simulation of social experience, *Perspectives on Psychological Science* 3 (2008) 173–192 (résumé vérifié sur la notice PubMed) : la fiction offre des « models or simulations of the social world via abstraction, simplification, and compression » et produit « a deep and immersive simulative experience of social interactions ». ↩
14. Elley/Mangubhai (chapitre 2, note 5 s.) ; sur la tradition de la lecture extensive, Day, R./Bamford, J., *Extensive Reading in the Second Language Classroom*, Cambridge 1998 (à vérifier sur l'imprimé) ; sur les bandes dessinées et la lecture facile, Krashen, S., *The Power of Reading*, 2e éd., Westport 2004, chap. 3 (à vérifier sur l'imprimé) ; Cary, S., *Going Graphic: Comics at Work in the Multilingual Classroom*, Portsmouth 2004 (à vérifier sur l'imprimé). ↩
15. Avertissement sur la transposition, en toute forme : sont documentées les fonctions de simulation et de compréhension des récits (notes 12 s.) et l'effet langagier d'une matière de lecture narrative (note 14). N'est pas documenté l'effet d'une bande dessinée déterminée comme toit narratif d'une plateforme d'apprentissage ; ce qui serait le plus aisément mesurable, c'est son usage (est-elle lue ? par qui ? avant ou après les films ?) — chapitre 9. ↩
16. La règle de protection est une appréciation pédagogique de ce livre, non un compte rendu de recherche ; elle découle de la sensibilité au traumatisme dans le travail avec des personnes réfugiées (cf. la littérature qualitative sur la community music, chapitre 6, note 12, qui souligne partout le volontariat et la propriété des contributions). ↩
17. Des lignes de recherche apparentées, mais non superposables, existent — le chant dans l'enseignement de la L2 (partie III), la community music avec des personnes réfugiées (chapitre 6, note 12), la traduction dans l'enseignement des langues (Translation in Language Teaching, parmi bien d'autres Cook, G., *Translation in Language Teaching*,

Oxford 2010, à vérifier sur l'imprimé) — ; l'intersection des trois est, selon la recherche menée ici, inexplorée. Si cela devait changer à l'issue d'une recherche bibliographique approfondie, le constat appartiendrait à la prochaine révision de ce livre. — Numérotation de cet appareil adaptée en rév. 3 (notes 9 s. nouvelles ; 9–14 anciennes → 11–16) ; en rév. 4 de nouveau : note fonctionnelle 11 nouvelle, 11–16 anciennes → 12–17. ↩

Partie VI — Les scénarios & la voie ouverte

> „Nicht die Wahrheit, in deren Besitz irgend ein Mensch ist, oder zu sein vermeinet, sondern die aufrichtige Mühe, die er angewandt hat, hinter die Wahrheit zu kommen, macht den Werth des Menschen."
> — Gotthold Ephraim Lessing, Eine Duplik (1778)

Chapitre 9 · Trois salles de classe et un programme honnête

9.1 Trois salles de classe

Tout ce que ce livre a développé converge dans trois scénarios concrets pour lesquels Pontivo est construit — et c'est la même machine qui porte les trois.

La première salle de classe se trouve en Allemagne : des personnes réfugiées apprennent l'allemand. Le film du soir — un documentaire en version originale allemande — tourne avec la piste cible allemande et, pour chacun, sa propre ligne en langue maternelle ; le clic explique le passage qui accroche ; une fois par semaine, le groupe chante, et à intervalles irréguliers une chanson venue d'Alep, de Kherson ou d'Asmara traverse l'atelier de traduction du chapitre 8. La deuxième salle de classe se trouve en Italie et ne diffère que par une position des axes : des personnes réfugiées apprennent l'italien — le chœur s'appelle ici *One Sky*, et c'est le scénario qui donne son nom à la bande dessinée. La troisième salle de classe est la classe de Bologne par laquelle ce livre a commencé : des adolescents italiens apprennent l'anglais — la même double piste, à cela près qu'en dessous se trouve maintenant l'italien et que, dans le lecteur, tourne le documentaire qui échouait jusqu'ici en cours sur la version doublée.

Que trois situations si différentes — intégration sous pression, intégration dans le pays d'arrivée, langue étrangère scolaire dans l'aisance — puissent être servies par un seul moteur n'est pas une pointe de marketing, mais une conséquence de la théorie : si l'acquisition d'une langue suit partout les mêmes conditions — input compréhensible, ancrage, anxiété basse —, alors les scénarios ne diffèrent que par la position des axes : quelle langue est cible, laquelle est échafaudage, qui conduit le groupe.[1]

Les trois salles de classe ne sont donc pas non plus une liste exhaustive. L'administration trie l'immigration en sphères séparées — ici les « expats », là les « personnes réfugiées » —, et dans les environnements d'apprentissage classiques, les groupes ne se rencontrent jamais. Sur le plan cognitif, pourtant, l'ingénieure irlandaise du bureau du groupe et le garçon syrien du chapitre 8 partagent la même expérience : la perte de la souveraineté langagière, l'épuisement du non-comprendre permanent, la tâche de se resituer. Pour les conditions de ce livre, le titre de séjour n'est pas une variable ; la même position des axes qui porte le cours d'intégration porte la spécialiste internationale et sa famille. Et elle porte l'école ordinaire, dont les classes sont depuis longtemps aussi plurilingues que n'importe quelle classe d'intégration : regarder et chanter ensemble dans des classes hétérogènes n'exige aucun seuil cognitif d'entrée — cela laisse les enfants agir comme groupe avant que l'échelle des notes ne les trie. Ce qui est né de la détresse des personnes réfugiées n'est, en aucun point, construit pour elles seules.

9.2 Amal, un an plus tard — une hypothèse

Il est tentant, à cet endroit, de dénouer la scène d'ouverture : Amal, après un an de films du soir et de répétitions de chœur, à la même caisse de supermarché — et cette fois elle répond. Ce livre ne cédera pas tout à fait à la tentation, car c'est précisément cette scène qui n'est pas un fait, mais l'*hypothèse* dont il s'agit ici d'examiner honnêtement la valeur. Ce que l'on peut dire est plus modeste et plus solide : après une telle année, Amal aurait entendu quelques centaines d'heures d'allemand véritable et compris — davantage de langue quotidienne parlée que six cents heures de cours ne peuvent jamais en contenir ; elle aurait des dizaines de tournures disponibles au mot près, parce qu'elle les a chantées ; et elle ne serait plus seule : des personnes à côté desquelles on a chanté pendant un an ne sont plus des inconnues devant lesquelles on se tait. Si cela suffit pour qu'elle comprenne „Sammeln Se Punkte?" et réponde non en riant, personne ne le sait — *encore* personne. La différence avec le système de cours du chapitre 1 n'est pas la promesse. La différence est que les conditions sont justes — et que l'on peut mesurer le résultat et que l'on veut le mesurer.

9.3 Ce que l'approche ne peut pas

Mais d'abord les limites, et les limites de fond, non les limites cosmétiques.

L'approche ne fait pas des locuteurs. La recherche qui donne à ce livre son fondement a un point faible, et il a été nommé au chapitre 2 : comprendre est nécessaire, mais vraisemblablement pas suffisant. Parler soi-même, en formulant librement — la négociation dans la conversation, la lutte pour sa propre phrase — a, d'après tout ce que montrent la recherche sur l'output et sur l'interaction, une fonction propre qu'aucun input ne remplace.[2] De la production, Pontivo ne couvre que le chant — une production littérale et protégée, précieuse et limitée. La conversation libre a lieu en dehors de la plateforme, ou pas du tout ; la tâche de la susciter reste aux enseignants, aux tandems, aux rencontres. Celui qui vendrait l'approche comme un substitut complet à l'enseignement des langues n'a pas lu ce livre.

L'approche présuppose la lecture. La double piste, la pièce maîtresse de la partie II, est du texte. Pour qui ne sait pas lire ou sait à peine lire — et dans les cours d'intégration sont assises des personnes qui n'ont jamais été alphabétisées ou qui apportent une autre écriture —, c'est précisément le mécanisme le mieux documenté ici qui ne porte pas.[3] Pour ces apprenants restent le son, l'image et la chanson ; si cela suffit est une question ouverte. Il serait facile de taire cette limite, puisqu'elle concerne le groupe le plus vulnérable.

L'approche dépend des droits. Un film authentique est la propriété d'autrui ; ce qui peut être montré, sous-titré et cliqué, ce sont les licences qui en décident, non les didacticiens. L'état des sources limite le catalogue, et un concept qui bâtit sur la quantité doit tenir cette

dépendance pour un risque d'exploitation — la section suivante déploie cette limite et les voies qui en sortent.

Le clic peut expliquer faux. La glose produite par la machine est une dette de qualité (chapitre 4) : sans vérification par échantillons contre une référence spécialisée, chaque explication isolée reste une petite promesse sans preuve.

Et l'enseignant reste irremplacé. Les tâches, la conversation, l'encouragement, la décision du moment où un échafaudage tombe — tout cela, la plateforme peut l'encadrer, mais non l'accomplir (chapitre 7). Là où elle serait mal comprise comme un modèle d'économie de personnel, elle serait pédagogiquement mal employée.

9.4 Le paradoxe des droits et le don de chansons

L'une des limites de 9.3 mérite une section propre, parce qu'elle n'est pas de nature cognitive mais juridique — et parce qu'on peut la façonner. Le droit d'auteur connaît certes une exception pour l'enseignement : le § 60a de la loi allemande sur le droit d'auteur (UrhG) permet aux établissements d'enseignement, à des fins non commerciales, d'utiliser jusqu'à 15 pour cent d'une œuvre publiée — les œuvres de faible étendue, et donc typiquement aussi des textes de chansons isolés, même intégralement. Mais l'exception a deux inconvénients, et tous deux frappent ce concept en son centre. Premièrement, elle privilégie les établissements d'enseignement eux-mêmes — écoles, universités, organismes de formation professionnelle et continue —, non une plateforme : pour son catalogue, Pontivo ne peut pas invoquer le § 60a ; n'est privilégié que son emploi *par* un établissement d'enseignement dans son propre enseignement. Deuxièmement, la loi exclut précisément les partitions : les enregistrements graphiques d'œuvres musicales ne peuvent pas être reproduits sans licence, même dans l'enseignement.[4] Mais une méthode qui vit de l'immersion dans des œuvres entières — le film entier, la chanson entière — ne va nulle part avec quinze pour cent. C'est le paradoxe des droits de l'éducation aux langues : précisément le matériau qui porte le plus, sur le plan langagier et émotionnel, est celui qu'il est le plus difficile d'utiliser avec sécurité juridique dans l'espace d'apprentissage.

La voie passe donc non par les exceptions, mais par les licences — et ainsi par le réseau international des sociétés de gestion collective : en Allemagne la GEMA et, pour les partitions et les textes de chansons, la VG Musikedition, en Autriche l'AKM, en Suisse la SUISA, à côté d'elles la SACEM, PRS, ASCAP et leurs sœurs. Leurs barèmes sont historiquement taillés pour la radiodiffusion, la représentation et le manuel scolaire, non pour des espaces d'apprentissage numériques protégés qui ont besoin d'œuvres entières.[5] Deux issues se dessinent. Les diffuseurs de service public et les diffuseurs financés par des fonds publics peuvent accorder des autorisations à des fins éducatives pour des environnements d'apprentissage fermés — leur mission le porte, et les données sur les sous-titres rassemblées au chapitre 3 en fournissent la

justification du même coup. Et pour les chansons, le concept établit un modèle de réciprocité, le *don de chansons* : des musiciennes, des musiciens et des chœurs utilisent le livre de chants et ses versions ChordPro comme outil de leur propre pratique de répétition et de scène ; en contrepartie, des titulaires de droits peuvent libérer certaines de leurs propres œuvres, par licence directe, pour l'espace didactique. Ainsi l'obstacle des droits devient un rapport d'échange — volontaire, documenté, résiliable. Si cela tient ne se décide pas au niveau du concept, mais à celui des contrats ; cela aussi fait partie du bilan honnête de ce chapitre.

9.5 La liste de mesure

Un concept qui se dit scientifique dit lui-même à quoi il pourrait échouer. La liste suivante est ordonnée selon la portée : en haut, les paris auxquels le concept est suspendu.

1. **La double piste.** Les débutants apprennent-ils de façon mesurable davantage avec une ligne cible et une ligne en langue maternelle synchrones qu'avec des sous-titres monolingues — sur des mois, non sur des minutes ? (Dispositif : groupes de comparaison sur le terrain ; tests de vocabulaire et de compréhension orale ; les constats sur extraits courts du chapitre 3 sont un point de départ, non une réponse.)

2. **L'effet du chœur sur la rétention.** Le module chœur abaisse-t-il l'anxiété langagière et augmente-t-il l'appartenance et la persévérance ? (Les instruments existent : l'échelle d'anxiété FLCAS, des mesures d'appartenance, tout simplement le taux d'abandon — la chaîne à deux étages du chapitre 6 devient ainsi vérifiable.[6])

3. **Le format de la chanson du pays d'origine.** Les mots issus de chansons traduites par les apprenants eux-mêmes restent-ils mieux que des mots de fréquence égale issus du corpus de films ? L'efficacité personnelle éprouvée de celles et ceux qui apportent la chanson augmente-t-elle ? (Petit dispositif avant/après par groupe de chœur ; le pari central du chapitre 8.)

4. **L'énoncé chanté.** Des tournures chantées sont-elles encore restituables après des semaines, là où des tournures entendues se sont effacées ? (Les constats de laboratoire du chapitre 5, transposés à l'échelle du terrain.)

5. **La glose au clic.** Audit par échantillons des explications machinales contre une référence spécialisée (taux d'erreur, types d'erreur) ; en outre : la rétention des mots cliqués par rapport aux mots non cliqués, et la question du flow — à partir de quelle densité de clics la soirée cinéma bascule-t-elle en travail de dictionnaire ?

6. **Gain de vocabulaire par heure de visionnage** comme grandeur de base courante (tests de vocabulaire par échantillons le long du matériau effectivement vu), afin de confronter à la réalité l'attente d'incidence du chapitre 4.

7. **La bande dessinée :** est-elle lue, par qui, et sa lecture précède-t-elle ou suit-elle l'usage des autres modules ? (Pure mesure d'usage — l'état des preuves n'autorise pas davantage, chapitre 8.)

Rien dans cette liste n'est exotique ; chaque ligne est réalisable avec les moyens du bord d'une plateforme et les instruments standard de la recherche. Ce qu'elle exige n'est qu'une seule chose : la volonté de se laisser mesurer — y compris au risque qu'un pari soit perdu. Des paris perdus ne sont pas une tare à l'aune de cette collection ; des paris non nommés en sont une.

9.6 L'arbre généalogique

Pour finir, l'inventaire que tout livre de cette collection se doit à lui-même : qu'y a-t-il là d'emprunté, qu'y a-t-il de propre ?

Emprunté, presque tout l'est, et à dessein. La perspective de l'acquisition plutôt que de l'apprentissage et les deux conditions viennent de Krashen et de la recherche qui l'a suivi ; l'architecture à deux canaux de la mémoire, de Paivio et Mayer ; les données sur les sous-titres, de Vanderplank à la méta-analyse de Louvain jusqu'aux études sur la double piste ; les constats sur les gloses et l'implication, de Nation, Laufer, Hulstijn et de la méta-analyse de Yanagisawa ; la mémoire du chant, de Wallace, Schön et Ludke, avec le désenchantement apporté par Racette et Peretz ; l'effet brise-glace, de Pearce, Launay et Dunbar ; l'échafaudage et la zone proximale de développement, de Vygotski, Wood, Bruner et Ross ; la réhabilitation de la langue maternelle, de Butzkamm et de la recherche sur le translanguaging ; la théorie du récit, de Bruner, Mar et Oatley ; la motivation et la sécurité, de Deci, Ryan, Csíkszentmihályi et Edmondson ; l'hypothèse du contact, d'Allport, Pettigrew et Tropp.[7]

Propres, trois choses le sont. Premièrement l'*intégration* : la réunion de ces fils dans une architecture — le film comme input, la chanson comme ancre, le groupe comme abaisseur de filtre — et dans un moteur qui sert les trois scénarios avec les mêmes axes. Deuxièmement le *format de la chanson du pays d'origine* comme exercice où les fils se croisent ; les éléments sont étrangers, la combinaison est, pour autant qu'on puisse en juger, nouvelle. Troisièmement l'*ancre mémorielle* comme grille d'examen : les trois questions avec lesquelles on peut interroger tout environnement d'apprentissage — y compris toute version future de Pontivo elle-même.

Et ainsi à la dernière application de cette grille, au livre lui-même. **Compréhensible ?** Il a construit sa ligne principale pour des lecteurs sans formation spécialisée et a déposé la profondeur dans les notes ; si c'est réussi, ce sont les cheffes de chœur et les enseignants qui le lisent qui en décident. **Ancré ?** Il a, là où il pouvait, raconté par des scènes plutôt que par des construits — par une femme à une caisse, par un garçon dont vingt voix chantent la chanson. **Relié ?** Il paraît ouvertement, lisible par tous, chez un éditeur qui ne met pas le savoir derrière

des barrières — et il invite expressément à la réfutation : chaque chiffre porte sa référence, chaque pari porte son nom.

Au commencement, il y avait une femme qui a payé en silence, bien qu'elle possède un certificat. Ce livre n'a pas promis qu'elle parlera. Il a montré sous quelles conditions les êtres humains commencent à parler — et il a décrit un instrument qui produit ces conditions, avec la liste à laquelle il doit se laisser examiner. Le reste n'est plus de la théorie. Le reste est un mardi soir dans une salle paroissiale, un film avec deux lignes de texte, et une chanson que quelqu'un a apportée.

Notes

1. L'architecture en axes est une description de produit (chapitre 7, note 11). La phrase « mêmes conditions, autre position des axes » est la conséquence de la théorie de l'acquisition développée dans la partie I et partage ses limites de validité — en particulier la limite de l'alphabétisation (9.3). ↩
2. Swain 1985 ; Long 1996 (tous deux chapitre 2, note 4, à vérifier sur l'imprimé). La délimitation suit le noyau consensuel marqué là : input nécessaire, output et interaction avec une fonction propre et irremplaçable. ↩
3. Sur l'éloignement de l'écrit et l'alphabétisation dans les cours d'intégration, le BAMF mène ses propres cours d'alphabétisation (cf. le rapport cité au chapitre 1, note 2, tableaux des types de cours) ; les données sur le captioning (chapitre 3) ont été obtenues partout sur des échantillons de personnes sachant lire. Pour des apprenants non alphabétisés, le mécanisme du sous-titre n'est pas documenté et est plausiblement sans effet — une limite du concept, non une note marginale. ↩
4. § 60a al. 1 UrhG (limite de 15 %, fins non commerciales, cercle d'utilisateurs privilégié), al. 2 (utilisation intégrale, entre autres, d'« œuvres de faible étendue » et d'œuvres indisponibles dans le commerce), al. 3 phrase 1 (entre autres, aucune « reproduction d'enregistrements graphiques d'œuvres musicales » = exclusion des partitions), al. 4 (établissements visés) — libellé vérifié sur le texte intégral officiel (gesetze-im-internet.de/urhg/__60a.html, consulté le 2026-07-25) ; introduit par la loi allemande sur le droit d'auteur dans la société de la connaissance (UrhWissG, 2017). Savoir si un texte de chanson concret est une « œuvre de faible étendue » se décide au cas par cas (accords-cadres conclus par les Länder avec les sociétés de gestion collective, à vérifier sur le libellé) ; l'obligation de rémunération au titre du § 60h UrhG reste intacte. Cette section ne remplace pas un conseil juridique dans le cas particulier. ↩
5. Sur la répartition des compétences (GEMA : droits de représentation et de communication ; VG Musikedition : entre autres les droits graphiques sur les partitions et les textes) et sur les sociétés sœurs internationales : présentations des sociétés elles-mêmes (à vérifier sur leurs barèmes et contrats respectifs). L'énoncé sur la taille des barèmes est une observation de structure, non une expertise tarifaire. ↩

6. FLCAS : Horwitz/Horwitz/Cope 1986 (chapitre 2, note 9), échelle p. 129 et s. ; mesures d'appartenance et de proximité sur le modèle de Pearce et al. 2015 (chapitre 6, note 1). Que la rétention (« retention ») convienne comme mesure d'intégration est une position de ce livre : qui reste reçoit de la quantité — et la quantité est, d'après la partie II, la première condition. ↩
7. L'ensemble des références se trouve aux chapitres 2 à 8 ; le registre des sources, en annexe, les rassemble, avec leur statut de vérification (vérifié / à vérifier sur l'imprimé) pour chaque référence — l'état de la vérification fait partie de la publication, non du matériau interne : cela aussi relève de « vérifiable plutôt qu'affirmé ». ↩

Annexe

Registre des sources

Alphabétique. Statut : **[V]** = vérifié à la source primaire (texte intégral ou résumé de l'éditeur ; chemin de vérification dans le registre des vérifications) · **[D]** = à vérifier sur l'imprimé ou le texte intégral (signalé comme tel dans la note correspondante). Références de chapitre entre parenthèses (« Kap. » = chapitre ; « Motto Teil » = exergue de partie). Le registre est conservé dans sa forme bibliographique originale dans toutes les éditions linguistiques.

- Allport, G. W., *The Nature of Prejudice*, Reading/MA 1954 — [D] (Kap. 8)
- Augustinus, *Confessiones*, Buch I, Kap. 14 (I,14,23) — [V, lateinischer Volltext] (Motto Teil II)
- BAMF, *Bericht zur Integrationskursgeschäftsstatistik für das Jahr 2024*, S. 17 ff., 19 — [V, PDF-Volltext] (Kap. 1, 9)
- Bruner, J., *Actual Minds, Possible Worlds*, Cambridge/MA 1986 — [D] (Kap. 8)
- Bruner, J., The narrative construction of reality, *Critical Inquiry* 18 (1991) 1–21 — [D] (Kap. 8)
- Butzkamm, W., *Psycholinguistik des Fremdsprachenunterrichts*, 3. Aufl. 2002 — [D] (Kap. 7)
- Butzkamm, W./Caldwell, J. A. W., *The Bilingual Reform*, Tübingen 2009 — [D] (Kap. 7)
- Cary, S., *Going Graphic: Comics at Work in the Multilingual Classroom*, Portsmouth 2004 — [D] (Kap. 8)
- Clift, S./Hancox, G., The perceived benefits of singing, *J Royal Society for the Promotion of Health* 121 (2001) 248–256 — [V, Abstract] (Kap. 6)
- Cook, G., *Translation in Language Teaching*, Oxford 2010 — [D] (Kap. 8)
- Council of Europe, *CEFR Companion Volume*, Straßburg 2020 — [V, Kernaussage an coe.int; Pinpoints am PDF zu prüfen] (Kap. 7)
- Csíkszentmihályi, M., *Flow: The Psychology of Optimal Experience*, New York 1990 — [D] (Kap. 2)
- Day, R./Bamford, J., *Extensive Reading in the Second Language Classroom*, Cambridge 1998 — [D] (Kap. 8)
- Deci, E. L./Ryan, R. M., *Intrinsic Motivation and Self-Determination in Human Behavior*, New York 1985 — [D] (Kap. 2)
- DeKeyser, R., Skill Acquisition Theory, in: VanPatten/Williams (Hg.), *Theories in SLA*, 2. Aufl. 2015 — [D] (Kap. 1)
- Dow, R./Vougioukalou, S., Wellbeing and Integration through Community Music, *Contemporary Music Review* 38 (2019) — [D, Verlagszugriff gesperrt] (Kap. 6)
- Dunbar, R., *Grooming, Gossip, and the Evolution of Language*, 1996 — [D] (Kap. 6)
- d'Ydewalle, G./Van de Poel, M., Incidental foreign-language acquisition by children watching subtitled TV, *J Psycholinguistic Research* 28 (1999) 227–244 — [D] (Kap. 3)
- Edmondson, A., Psychological Safety and Learning Behavior in Work Teams, *Administrative Science Quarterly* 44 (1999) 350–383 — [V, Definition am Volltext-

Spiegel; am Druck zu prüfen] (Kap. 6)
- Elley, W./Cutting, B./Mangubhai, F./Hugo, C., *Lifting Literacy Levels with Story Books*, ERIC ED416441 (1996), S. 4 f. — [V, Volltext] (Kap. 2, 7, 8)
- Elley, W./Mangubhai, F., The impact of reading on second language learning, *Reading Research Quarterly* 19 (1983) 53–67 — [D; Befunde nach der Autoren-Darstellung ED416441 referiert] (Kap. 2)
- Ellis, R., *Task-based Language Learning and Teaching*, Oxford 2003 — [D] (Kap. 7)
- Ellis, R., *The Study of Second Language Acquisition*, 2. Aufl., Oxford 2008 — [D] (Kap. 1)
- Europarat, *Gemeinsamer europäischer Referenzrahmen für Sprachen* (2001), Globalskala — [D; Raster online geprüft] (Kap. 1)
- García, O., *Bilingual Education in the 21st Century*, Malden 2009 — [D] (Kap. 7)
- García, O./Wei, L., *Translanguaging*, Basingstoke 2014 — [D] (Kap. 7, 8)
- Hattie, J., *Visible Learning*, London 2009 — [D; Effektstärken tw. sekundär, am Druck zu prüfen] (Kap. 7)
- Herder, J. G., Von Ähnlichkeit der mittlern englischen und deutschen Dichtkunst, *Deutsches Museum* 1777, S. 433 — [V, Digitalisat] (Motto Teil V)
- Horwitz, E. K./Horwitz, M. B./Cope, J., Foreign Language Classroom Anxiety, *Modern Language Journal* 70 (1986) 125–132 — [V, Volltext; Definition S. 128, FLCAS S. 129 f.] (Kap. 2, 6, 8, 9)
- Hulstijn, J., Incidental and intentional learning, in: Doughty/Long (Hg.), *Handbook of SLA*, Oxford 2003, S. 349–381 — [D] (Kap. 4)
- INVALSI, Ergebnisdarstellung Prova INVALSI 2025, Inglese grado 13 (invalsi.it) — [V, Institutsseite; Rapporto Nazionale am Volltext zu prüfen] (Kap. 1)
- IntV (Integrationskursverordnung), § 10 Abs. 1, § 11 Abs. 1 — [V, gesetze-im-internet.de] (Kap. 1)
- Kemmis, S./McTaggart, R./Nixon, R., *The Action Research Planner*, Singapur 2014 — [D] (Kap. 7)
- Koelsch, S., *Brain and Music*, Chichester 2012 — [D] (Kap. 8)
- Krashen, S., *Principles and Practice in Second Language Acquisition*, Pergamon 1982 (Internet-Ausgabe 2009), S. 10 f., 20 f., 30–32 — [V, Volltext] (Kap. 1, 2)
- Krashen, S., *The Input Hypothesis*, London 1985 — [D] (Kap. 2)
- Krashen, S., *The Power of Reading*, 2. Aufl., Westport 2004 — [D] (Kap. 8)
- Kreutz, G. et al., Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state, *J Behavioral Medicine* 27 (2004) 623–635 — [V, Abstract] (Kap. 6)
- Laufer, B./Hulstijn, J., Incidental vocabulary acquisition: The construct of task-induced involvement, *Applied Linguistics* 22 (2001) 1–26 — [D] (Kap. 4, 8)
- Lessing, G. E., *Eine Duplik* (1778), Erster Abschnitt — [V, Wortlaut an geprüften Zitatquellen; am Druck zu prüfen] (Motto Teil VI)

- Liao, S./Kruger, J.-L./Doherty, S., The impact of monolingual and bilingual subtitles, *J Specialised Translation* 33 (2020) 70–98 — [V, Open-Access-Volltext] (Kap. 3, 7)
- Lightbown, P./Spada, N., *How Languages are Learned*, 4. Aufl., Oxford 2013 — [D] (Kap. 1)
- Long, M., The role of the linguistic environment in SLA, in: Ritchie/Bhatia (Hg.), *Handbook of SLA*, 1996, S. 413–468 — [D] (Kap. 2, 9)
- Long, M., *SLA and Task-Based Language Teaching*, Malden 2015 — [D] (Kap. 7)
- Ludke, K. M./Ferreira, F./Overy, K., Singing can facilitate foreign language learning, *Memory & Cognition* 42 (2014) 41–52 — [V, Volltext] (Kap. 5, 8)
- Mar, R. A./Oatley, K., The function of fiction is the abstraction and simulation of social experience, *Perspectives on Psychological Science* 3 (2008) 173–192 — [V, Abstract] (Kap. 8)
- Mayer, R. E., *Multimedia Learning*, 3. Aufl., Cambridge 2021 — [D] (Kap. 3)
- McLaughlin, B., *Theories of Second-Language Learning*, London 1987 — [D] (Kap. 1, 2)
- Montero Perez, M./Peters, E./Clarebout, G./Desmet, P., Effects of captioning on video comprehension and incidental vocabulary learning, *Language Learning & Technology* 18/1 (2014) 118–141 — [V, Volltext] (Kap. 4)
- Montero Perez, M./Van Den Noortgate, W./Desmet, P., Captioned video for L2 listening and vocabulary learning: A meta-analysis, *System* 41 (2013) 720–739 — [V, Abstract-Kernaussage; g-Werte am Volltext zu prüfen] (Kap. 3)
- Murphey, T., The song stuck in my head phenomenon, *System* 18 (1990) 53–64 — [D] (Kap. 5)
- Murphey, T., *Music and Song*, Oxford 1992 — [D] (Kap. 5)
- Nation, I. S. P., *Learning Vocabulary in Another Language*, 2. Aufl., Cambridge 2013 — [D] (Kap. 4)
- Nietzsche, F., *Götzen-Dämmerung*, Sprüche und Pfeile 33 (1889) — [V, Gutenberg-Volltext] (Motto Teil III)
- Norris, J./Ortega, L., Effectiveness of L2 instruction, *Language Learning* 50 (2000) — [D] (Kap. 1)
- Oldenburg, R., *The Great Good Place*, New York 1989 — [D] (Kap. 6)
- Paivio, A., *Mental Representations: A Dual Coding Approach*, Oxford 1986 — [D] (Kap. 3)
- Paivio, A., Dual coding theory: Retrospect and current status, *Canadian J Psychology* 45 (1991) 255–287 — [D] (Kap. 3)
- Patel, A., *Music, Language, and the Brain*, Oxford 2008 — [D] (Kap. 5)
- Pearce, E./Launay, J./Dunbar, R. I. M., The ice-breaker effect, *Royal Society Open Science* 2 (2015) 150221 — [V, Abstract wörtlich] (Kap. 6, 9)
- Pettigrew, T. F./Tropp, L. R., A meta-analytic test of intergroup contact theory, *J Personality and Social Psychology* 90 (2006) 751–783 — [V, Abstract] (Kap. 8)
- Racette, A./Peretz, I., Learning lyrics: to sing or not to sing?, *Memory & Cognition* 35 (2007) 242–253 — [V, Abstract] (Kap. 5, 8)

- Richards, J. C./Rodgers, T. S., *Approaches and Methods in Language Teaching*, 3. Aufl., Cambridge 2014 — [D] (Kap. 7)
- Rosa, H., *Resonanz. Eine Soziologie der Weltbeziehung*, Berlin 2016 — [D] (Kap. 8)
- Ryan, R. M./Deci, E. L., Self-determination theory and the facilitation of intrinsic motivation, *American Psychologist* 55 (2000) 68–78 — [D] (Kap. 2)
- Rupérez Micola, A./Aparicio Fenoll, A./Banal-Estañol, A./Bris, A., TV or not TV? The impact of subtitling on English skills, *J Economic Behavior & Organization* 158 (2019) 487–499 — [V, Volltext] (Kap. 3)
- Schön, D. et al., Songs as an aid for language acquisition, *Cognition* 106 (2008) 975–983 — [V, Abstract; Experimentdetails am Volltext zu prüfen] (Kap. 5)
- Seume, J. G., *Die Gesänge* (1804) — [V, Digitalisate] (Motto Teil IV)
- Slamecka, N. J./Graf, P., The generation effect, *JEP: Human Learning and Memory* 4 (1978) 592–604 — [V, Abstract] (Kap. 8)
- Streifler, D., *Kalibriertes Vertrauen*, ra.de-Verlag 2026 — [V, Schwesterband der Reihe] (Kap. 4)
- Swain, M., Communicative competence, in: Gass/Madden (Hg.), *Input in SLA*, Rowley 1985, S. 235–253 — [D] (Kap. 2, 9)
- Teimouri, Y./Goetze, J./Plonsky, L., Second language anxiety and achievement: A meta-analysis, *SSLA* 41 (2019) 363–387 — [D] (Kap. 2)
- Twain, M., *A Tramp Abroad*, Anhang D: The Awful German Language (1880) — [V, Wikisource-Transkription] (Motto Teil I)
- UrhG (Urheberrechtsgesetz), § 60a (mit § 60h) — [V, gesetze-im-internet.de, abgerufen 2026-07-25] (Kap. 9)
- Vanderplank, R., The value of teletext sub-titles in language learning, *ELT Journal* 42 (1988) 272–281 — [D] (Kap. 3)
- Vanderplank, R., *Captioned Media in Foreign Language Learning and Teaching*, London 2016 — [D] (Kap. 3)
- Vygotskij, L. S., *Mind in Society*, Cambridge/MA 1978, S. 86 — [V, Standardrezeption mehrfach; am Druck zu prüfen] (Kap. 7)
- Wallace, W. T., Memory for music: Effect of melody on recall of text, *JEP: LMC* 20 (1994) 1471–1485 — [V, Abstract/Rezeption; am Volltext zu prüfen] (Kap. 5, 8)
- Wang, A./Pellicer-Sánchez, A., Incidental vocabulary learning from bilingual subtitled viewing, *Language Learning* 72 (2022) — [D] (Kap. 3)
- Welch, G. F. et al., Singing and social inclusion, *Frontiers in Psychology* 5 (2014) 803 — [V, PMC-Volltext] (Kap. 6, 8)
- Winke, P./Gass, S./Sydorenko, T., The effects of captioning videos, *Language Learning & Technology* 14/1 (2010) 65–86 — [V, Volltext] (Kap. 3)
- Wood, D./Bruner, J. S./Ross, G., The role of tutoring in problem solving, *J Child Psychology and Psychiatry* 17 (1976) 89–100 — [D, Verlagszugriff gesperrt] (Kap. 7)
- Yanagisawa, A./Webb, S./Uchihara, T., How do different forms of glossing contribute to L2 vocabulary learning from reading?, *SSLA* 42 (2020) 411–438 — [V, Verlagsabstract]

(Kap. 4, 7, 8)

Abréviations

- **BAMF** — Bundesamt für Migration und Flüchtlinge (Office fédéral allemand de la migration et des réfugiés)
- **DTZ** — Deutsch-Test für Zuwanderer (test d'allemand pour les immigrés)
- **ES** — taille d'effet
- **FLCAS** — Foreign Language Classroom Anxiety Scale
- **CECR/GER** — Cadre européen commun de référence pour les langues
- **IntV** — Integrationskursverordnung (règlement allemand sur les cours d'intégration)
- **INVALSI** — Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione (institut national italien d'évaluation du système éducatif)
- **JEP** — Journal of Experimental Psychology
- **L1 / L2** — première langue (langue maternelle) / langue seconde ou étrangère
- **LLT** — Language Learning & Technology (revue)
- **SLA** — Second Language Acquisition (acquisition d'une langue seconde)
- **SSLA** — Studies in Second Language Acquisition (revue)
- **TBLT** — Task-Based Language Teaching (enseignement des langues par les tâches)
- **TOEFL** — Test of English as a Foreign Language
- **UE** — Unterrichtseinheit(en) (unités ou heures de cours)
- **ZPD** — Zone of Proximal Development (zone proximale de développement)

Glossaire

- **Ancre mémorielle** — la grille d'examen de ce livre : « Compréhensible ? Ancré ? Relié ? » (chap. 2).
- **Apprentissage incident** — apprentissage comme produit dérivé d'une activité de compréhension, sans intention d'apprendre (chap. 4).
- **Captions / sous-titres en langue cible** — sous-titres dans la même langue que le son ; ils segmentent le flux sonore et relient le son à l'écrit (chap. 3).
- **CECR** — Cadre européen commun de référence pour les langues ; niveaux A1–C2, B1 = « niveau seuil » (chap. 1).
- **Charge d'implication (involvement load)** — le modèle à trois facteurs de la profondeur de traitement dans l'apprentissage des mots, de Laufer et Hulstijn : besoin, recherche, évaluation (chap. 4).
- **ChordPro** — format texte pour les partitions de chansons, qui intègre les symboles d'accords dans les paroles ; standard de la pratique du chant en commun, base du module livre de chants (chap. 5).
- **Don de chansons** — le modèle de réciprocité de ce concept : usage du livre de chants pour la scène et les répétitions en échange d'une licence directe sur des œuvres isolées pour l'espace d'apprentissage (chap. 9).
- **Double codage (dual coding)** — la théorie de Paivio de deux systèmes de mémoire reliés (verbal et imagé) ; un contenu doublement codé a davantage de chemins de rappel (chap. 3).
- **Doubles sous-titres (sous-titres bilingues)** — affichage simultané de la ligne en langue cible et de la ligne en langue maternelle ; la forme construite du lecteur de Pontivo (chap. 3).
- **DTZ** — « Deutsch-Test für Zuwanderer », test final échelonné du cours d'intégration allemand (B1/A2) (chap. 1).
- **Effet de génération** — un contenu produit par soi-même est mieux retenu qu'un contenu seulement lu (Slamecka & Graf 1978) (chap. 8).
- **Étayage (scaffolding)** — structure de soutien démontable qu'une personne plus capable érige pour un apprenant (Wood/Bruner/Ross 1976) (chap. 7).
- **Filtre affectif** — l'image de Krashen pour le blocage entre l'input compris et l'acquisition de la langue, dû à l'anxiété, au manque de confiance en soi ou au manque de motivation ; filtre haut = peu de choses arrivent (chap. 2).
- **FLCAS** — Foreign Language Classroom Anxiety Scale ; échelle standard de mesure de l'anxiété langagière (Horwitz et al. 1986) (chap. 2, 9).
- **Flow** — l'état d'activité auto-portée décrit par Csíkszentmihályi, dans lequel l'effort disparaît du vécu (chap. 2).
- **Glose** — brève explication de mot en marge d'un texte ; sous forme numérique : l'explication au clic (chap. 4).

- **Hypothèse du contact** — les conditions d'Allport sous lesquelles la rencontre entre groupes défait les préjugés (égalité de statut, but commun) ; confirmée de façon méta-analytique par Pettigrew/Tropp (chap. 8).
- **i + 1** — la formule de Krashen pour un input juste au-dessus du niveau actuel *i* ; l'acquisition se produit en comprenant un tel input (chap. 2).
- **Misgrounding** — une référence qui existe, mais qui ne porte pas l'énoncé (terme issu du volume frère *Kalibriertes Vertrauen*) (chap. 4).
- **Paradoxe des droits** — constat de ce livre : le matériau le plus efficace sur le plan langagier et émotionnel est celui qu'il est le plus difficile d'utiliser avec sécurité juridique dans l'espace d'apprentissage (chap. 9.4).
- **Sécurité psychologique** — chez Edmondson, la conviction partagée d'un groupe que le risque interpersonnel (demander, se tromper) n'est pas sanctionné ; condition préalable de l'apprentissage (chap. 6).
- **Théorie de l'autodétermination (SDT)** — la théorie de la motivation de Deci et Ryan : autonomie, sentiment de compétence et affiliation sociale comme besoins fondamentaux à partir desquels l'engagement se porte lui-même (chap. 2).
- **Tiers-lieu** — terme d'Oldenburg pour les espaces communautaires neutres et niveleurs de statut, à côté du domicile et du travail ; dans ce livre : le chœur (chap. 6).
- **Translanguaging** — compréhension du plurilinguisme comme un répertoire commun ; sur le plan pédagogique : la première langue comme outil légitime en classe (chap. 7).
- **Zone proximale de développement (ZPD)** — l'intervalle de Vygotski entre ce qu'un apprenant sait faire seul et ce qu'il sait faire avec de l'aide ; le lieu de l'apprentissage (chap. 7).

Sur le concept & l'équipe

Pontivo (pontivo.de) est une plateforme d'acquisition des langues par le film, la chanson et la communauté : de la vidéo authentique avec un lecteur de sous-titres bilingue et synchrone et une explication au clic ; un livre de chants avec accords pour le chant commun en groupe, doté d'un mode enseignant/élève ; une salle de classe pour des scénarios d'apprentissage reposant sur le même moteur plurilingue ; et la bande dessinée « Coro One Sky » comme toit narratif. La plateforme est développée par une petite équipe autour de la ra.de Verlags-KG.

Ce livre est à Pontivo ce que le volume frère *Kalibriertes Vertrauen* est à trust-on.net : la justification scientifique ouverte de l'approche — y compris la liste de ce qui n'est pas encore prouvé. Il paraît comme volume 1 de la collection **« Schriften zur Sprachbildung mit Film und Musik »** (Écrits sur l'éducation aux langues par le film et la musique) aux éditions ra.de (Diamond Open Access : librement lisible, sans frais ni pour les lecteurs ni pour les auteurs).

Vérifiabilité. Chaque référence de ce livre porte un statut de vérification (vérifiée à la source primaire, ou expressément « à vérifier sur l'imprimé ») ; le registre des vérifications fait partie des matériaux de publication. Les corrections et les réfutations sont bienvenues et sont intégrées dans les révisions avec un protocole des modifications — l'aune de la collection vaut aussi pour ce livre : vérifiable plutôt qu'affirmé.

Ce texte français est une traduction de l'original allemand ; en cas de doute, la version allemande fait foi.