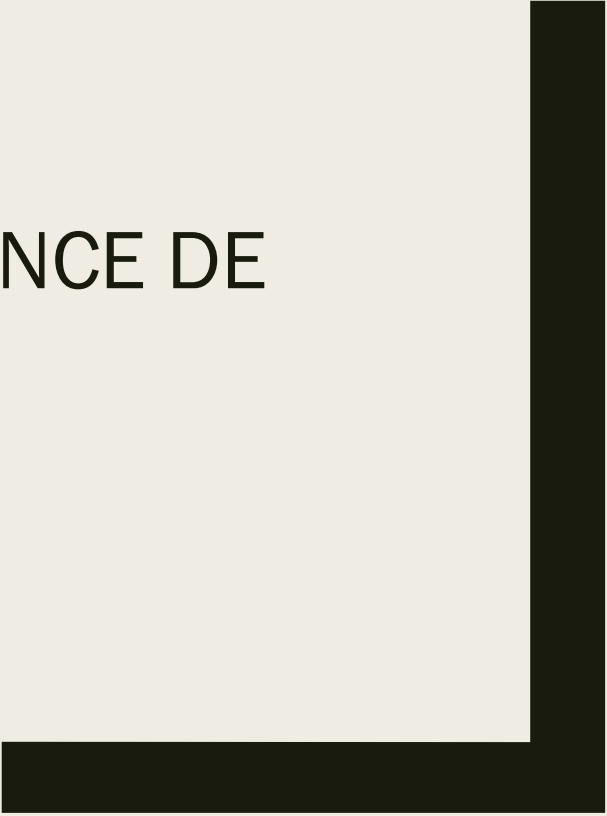




MODÈLES DE RECONNAISSANCE DE MOTS BILINGUE

Séverine Casalis
MIASHS, modélisation L3
2021 2022



Les modèles de RM mots bilingue

- Les questions
- 1 ou 2 lexique
- Accès sélectif ou non sélectif
- Le modèle de base

Le modèle BIA

- Bilingual Interactive Activation
- Dijkstra & Van Heuven, 1998

Base théorique; le modèle IA (Rumelhart & McClelland, 1981)

- Cf cours « psychologie générale»
- Cf ouvrages Ferrand (Ferrand, L. (2001). *Cognition et Lecture. Processus de base de la reconnaissance des mots écrits*. Bruxelles : De Boeck Université.
- Ferrand, L. (2007). [*Psychologie cognitive de la lecture*](#). Bruxelles : De Boeck Université, Collection Ouvertures Psychologiques. ou Ferrand, L., & Ayora, P. (2009). [*Abrégé de Psychologie Cognitive de la Lecture*](#). Bruxelles : De Boeck Université, Collection Préparer l'examen.)
- Les trois niveaux d'activation: traits, lettres, mots
- Facilitation inter niveaux, inhibition (essentiellement) intra niveau. Le mécanisme d'inhibition entre les mots (compétition lexicale) relève l'existence de relations entre les mots au sein du lexique.

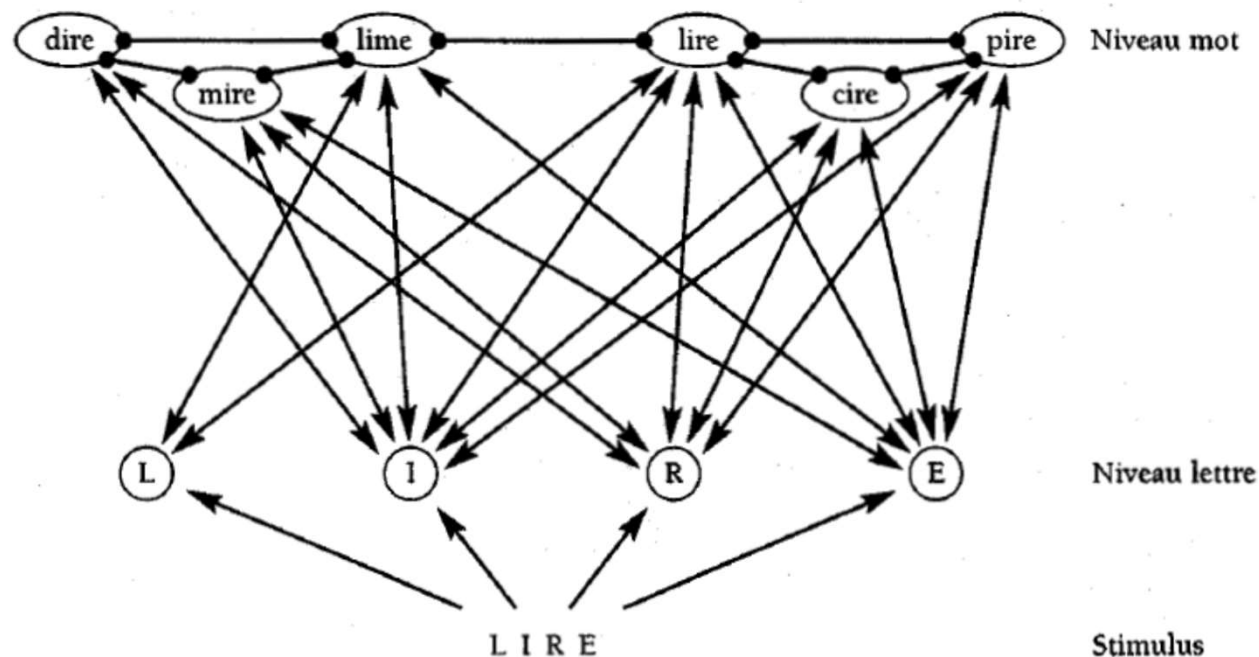


Figure 1. – Le modèle d'activation-interactive (McClelland & Rumelhart, 1981). Les flèches représentent des connexions de nature excitatrice entre unités de traitement, tandis que les points représentent des inhibitions (seules les connexions inhibitrices entre les représentations des mots figurent ici). Sur présentation d'un mot (ex. : lire), les représentations des lettres du mot sont activées, ce qui provoque l'activation de toutes les représentations des mots qui ont au moins une lettre en commun avec le stimulus. Les représentations des mots les plus activées sont celles qui correspondent aux voisins orthographiques du mot stimulus (ex. : dire, mire, lime, cire, pire). Toutes ces représentations s'inhibent mutuellement.

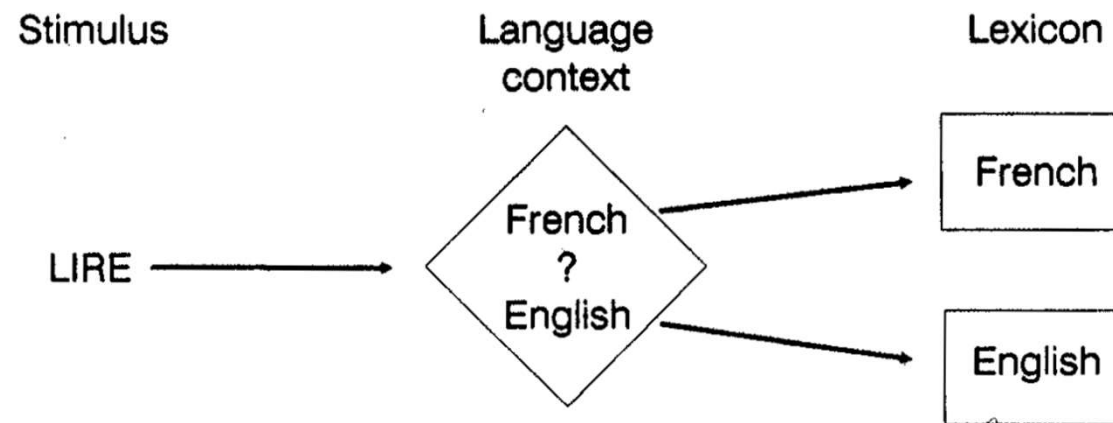


Figure 1. Representation of the language-selective access hypothesis according to which incoming sensory information is guided to one of the other of the bilingual's lexicons by language context information

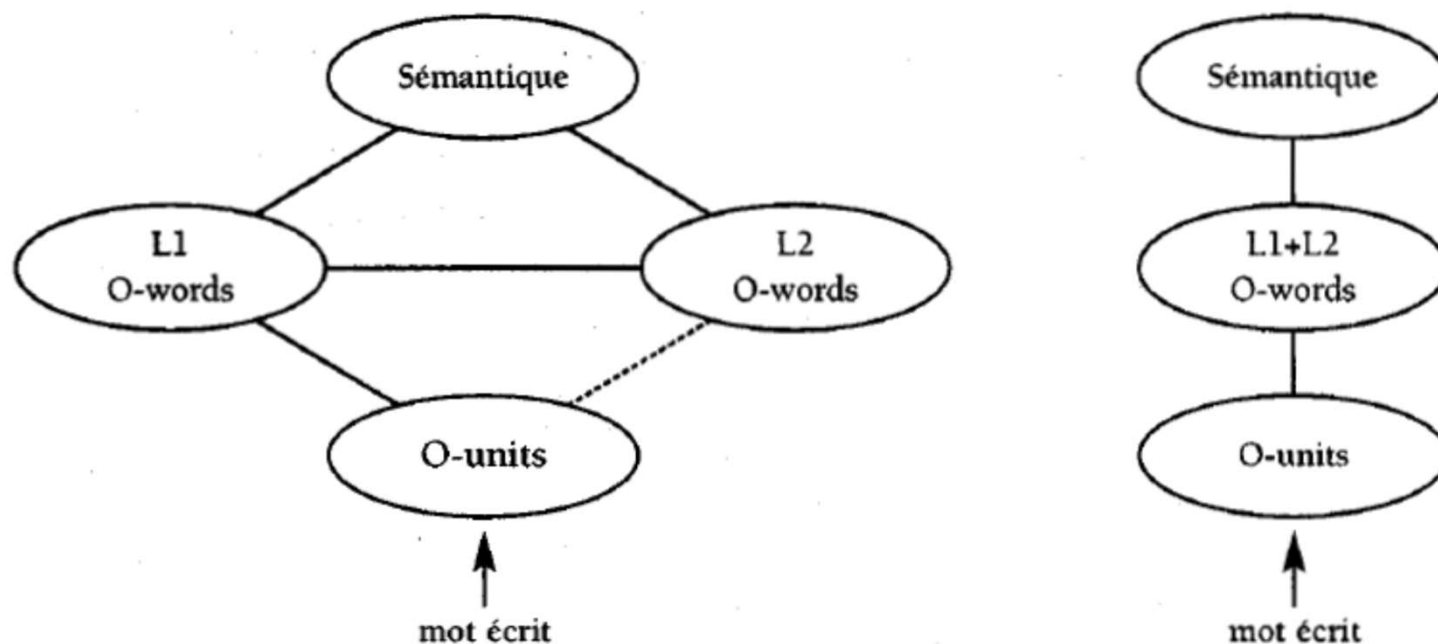


Figure 7. – Deux architectures possibles pour le codage orthographique en situation bilingue. À gauche, un système avec lexique orthographique indépendant pour chaque langue, et la possibilité de diriger le signal vers l'un ou l'autre de ces lexiques. À droite, un système avec un seul lexique orthographique intégré pour les deux langues.

Combien de lexique(s) pour le bilingue?

Lexique	Lexiques séparés	Lexique intégré
Accès	Sélectif	Non Sélectif

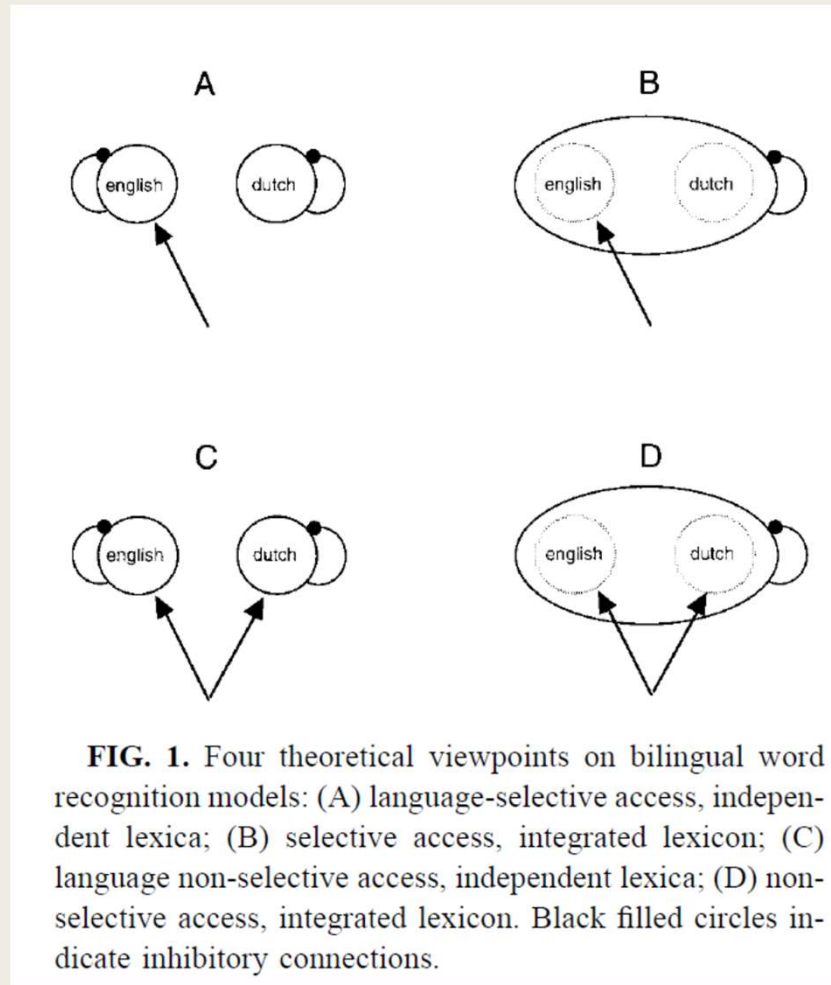
Lexique	Lexiques séparés	Lexique intégré
Accès	Sélectif	Non Sélectif

Quelques observations

- Faits marquants en production
 - *pas d'interférence visible chez les bilingues*
 - *: accent + emprunt à la syntaxe*
- Dans domaine de la compréhension, dépend de la façon dont les mots stockés
 - *Soit deux lexiques*
 - *Soit un seul lexique intégré*

Les hypothèses lexiques intégré/séparé

Acces sélectif/ non sélectif



Les hypothèses lexiques intégré/séparé

Acces sélectif/ non sélectif

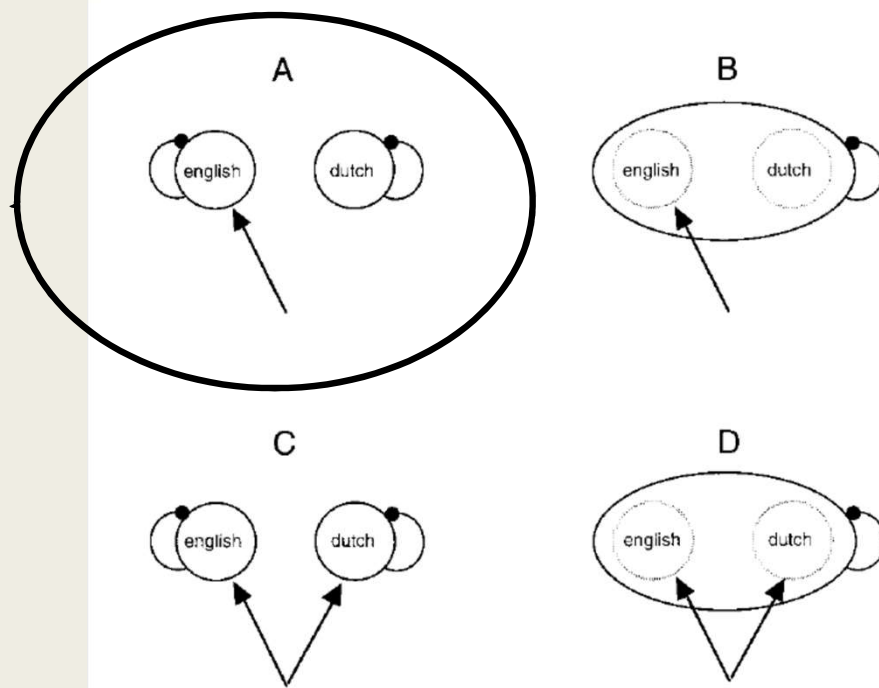
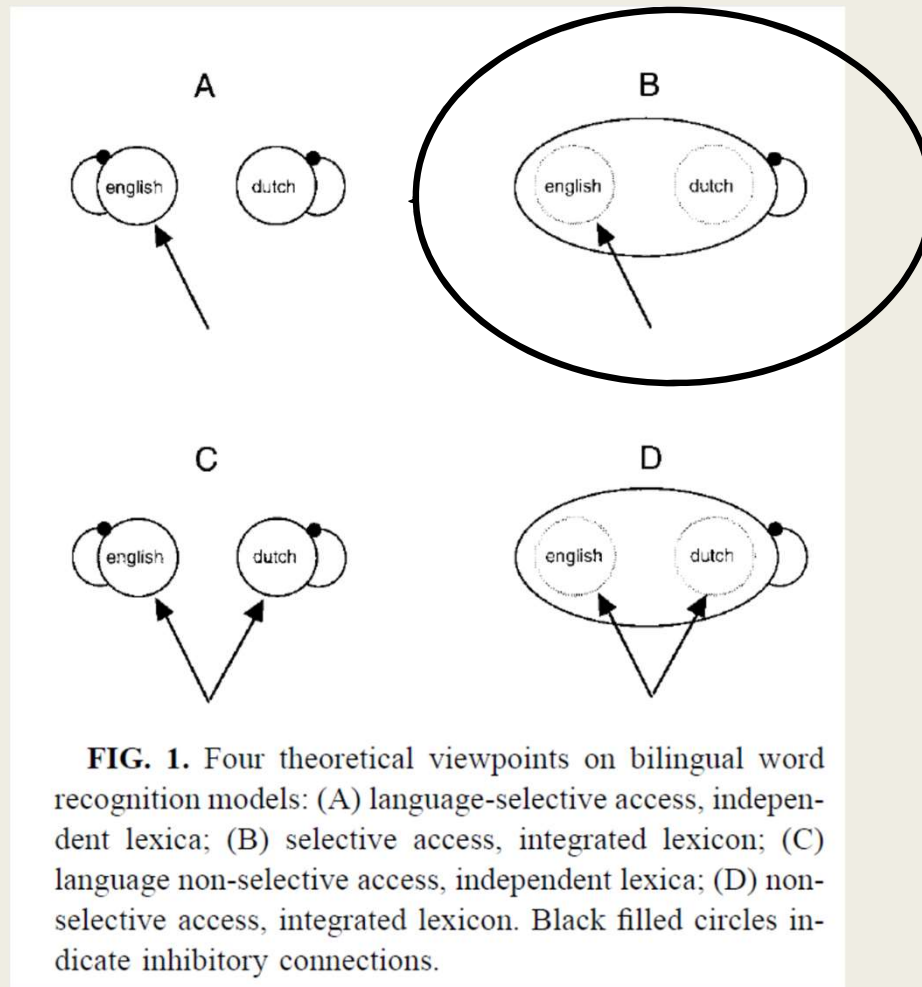


FIG. 1. Four theoretical viewpoints on bilingual word recognition models: (A) language-selective access, independent lexica; (B) selective access, integrated lexicon; (C) language non-selective access, independent lexica; (D) non-selective access, integrated lexicon. Black filled circles indicate inhibitory connections.

Note: en cas de lexiques séparés, pas d'effets inhibiteurs latéraux

Les hypothèses lexiques intégré/séparé

Acces sélectif/ non sélectif



Note: en cas de lexiques séparés, pas d'effets inhibiteurs latéraux

Les hypothèses lexiques intégré/séparé

Accès

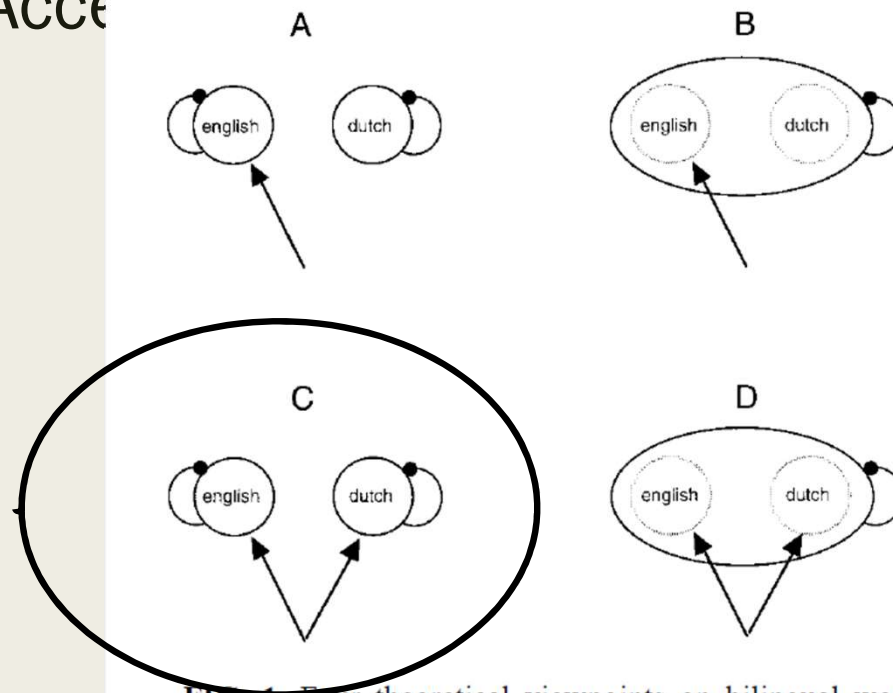
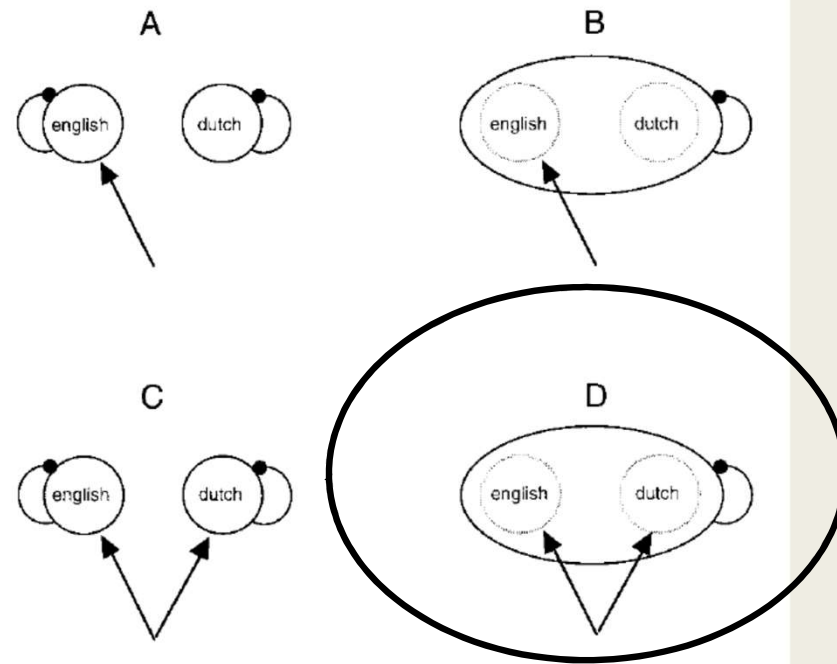


FIG. 1. Four theoretical viewpoints on bilingual word recognition models: (A) language-selective access, independent lexica; (B) selective access, integrated lexicon; (C) language non-selective access, independent lexica; (D) non-selective access, integrated lexicon. Black filled circles indicate inhibitory connections.

Note: en cas de lexiques séparés, pas d'effets inhibiteurs latéraux

Les hypothèses lexiques intégré/séparé

Accès sélectif/ non sélectif



Note: en cas de lexiques séparés, pas d'effets inhibiteurs latéraux

FIG. 1. Four theoretical viewpoints on bilingual word recognition models: (A) language-selective access, independent lexica; (B) selective access, integrated lexicon; (C) language non-selective access, independent lexica; (D) non-selective access, integrated lexicon. Black filled circles indicate inhibitory connections.

Premières données sur sélectivité

Résultats en faveur sélectivité

Macnamara and Kushnir (1971) bilingues doivent juger de valeur vérité de phrases écrites et parlées dans langues mixées : temps plus long dans phrases bilingues que monolingues

Soares and Grosjean (1984) :

- bilingues rejettent plus lentement pseudomots que monolingues

- bilingues identifient plus lentements mots en conditions code switching que monolingue

Conclusion : selection de langue a priori

Scarborough, Gerard, and Cortese (1984), bilingues Anglais / espagnols réponses aux mots non cibles (espagnols) comme si pseudomots (même Vitesse et pas d'effet de fréquence)

Données en faveur de la non-sélectivité

Dyer, 1971 : the Stroop color–word task

Ehri & Bouchard-Ryan, 1980 : picture–word interference task

Flanker >>> interference

Dans les tâches de décision lexicale :

En DL allemande Le temps pour rejeter pseudomots affecté par la “légalité” des pseudomots mots aussi bien en allemand (PFLOK) qu’en anglais (TWOUL) Altenberg and Cairns (1983)

- Nas (1983) Dutch/English bilinguals mettent plus de temps à rejeter PM semblables à mots dutch en DL anglaise (
- Guttentag, Haith, & Goodman, 1984 : interlangue de même amplitude que effets intralangue ortho ou phono) similarité

Comment travailler sur les réponses positives?

exploiter l'effet du voisinage?

Les expériences conduites par VanHeuven et al 1998

Stimuli

Une liste de 1323 mots anglais et 982 mots néerlandais a d'abord été générée.

Pour chaque mot: calcul du nombre de voisins dans chaque langue, et de leur fréquence

- Tous les items monolingues (pas d'homographe pas de cognate)

		Nb voisins langue cible	
		Beaucoup	peu
Nb voisins langue non cible	Beaucoup	20 items	20 items
	peu	20 items	20 items

Méthode utilisée: démasquage progressif

- Alternance de présentation du mot cible et du masque
- Temps de présentation du mot cible augmente progressivement (temps masque baisse progressivement)
 - *15 ms pour le mot, 285 pour le masque*
 - *Pas de 15 ms*
- Le participant appuie quand il a reconnu le mot
- On relève le temps de réponse

2 groupes de participants

- Néerlandais Très compétents en anglais
- Néerlandais moins compétents en anglais

TABLE 2

Mean Identification Latencies of Experiment 1 (in Milliseconds) of Participants Presented with Dutch Words in the First Block or the Second Block (Standard Deviation and Percentage of Errors Are Presented in Parentheses.)

Dutch N	English N	Low proficiency		High proficiency		Mean
		Block 1	Block 2	Block 1	Block 2	
Large	large	1681 (310, 2.7)	1659 (182, 3.5)	1739 (166, 2.7)	1674 (219, 3.5)	1689
Large	small	1684 (329, 1.8)	1634 (170, 3.0)	1675 (229, 3.6)	1672 (180, 5.5)	1667
Small	large	1614 (275, 1.8)	1600 (179, 3.0)	1707 (163, 1.4)	1595 (186, 4.5)	1630
Small	small	1586 (258, 4.5)	1529 (135, 1.5)	1639 (217, 3.2)	1599 (217, 2.5)	1589

- Conclusion: le temps de reconnaissance de mots dépend de la taille du voisinage de la langue non cible
 - *Activation des 2 langues*
 - *Cependant rôle compétence pas très clair*
 - *Question à poser: les plus compétents ont un meilleur contrôle sur les langues*

Expé 2: idem mais en présentation mixée

TABLE 4

Mean Identification Latencies of Experiment 2 (in Milliseconds) (Standard Deviation and Percentage of Errors Are Presented in Parentheses.)

Dutch N	English N	English	Dutch
Large	large	1777 (247, 4.8)	1756 (249, 7.0)
Large	small	1761 (245, 5.5)	1710 (205, 5.1)
Small	large	1660 (233, 3.1)	1689 (229, 4.4)
Small	small	1739 (240, 3.0)	1637 (189, 4.0)

- Effet inhibition liée à voisinage autre langue
- Effet plus fort en présentation mixee plutôt qu'ne blocs

Expérience 3: réplication avec autre paradigme: la décision lexicale

TABLE 5

Mean Identification Latencies of Experiment 3 (in Milliseconds) (Standard Deviation and Percentage of Errors Are Presented in Parentheses.)

Dutch N	English N	English	Dutch	Nonwords
Large	large	621 (69, 13.3)	581 (59, 7.6)	675 (75, 9.7)
Large	small	642 (85, 12.6)	566 (60, 7.5)	672 (78, 9.7)
Small	large	590 (62, 6.6)	561 (58, 6.6)	647 (79, 6.7)
Small	small	619 (73, 8.7)	560 (56, 6.4)	649 (73, 6.4)

Expé 4: mots anglais en L1 (monolingues) et L2 (bilingues)

TABLE 6

Mean Identification Latencies of Experiment 4 (in Milliseconds) of Monolingual and Bilingual Participants
(Standard Deviation and Percentage of Errors Are Presented in Parentheses.)

Dutch N	English N	Monolinguals		Bilinguals	
		Words	Nonwords	Words	Nonwords
Large	large	488 (58, 4.0)	575 (86, 8.8)	585 (69, 12.1)	651 (94, 9.5)
Large	small	507 (71, 5.0)	558 (74, 4.5)	583 (74, 12.1)	635 (94, 4.0)
Small	large	484 (57, 3.0)	581 (87, 5.8)	561 (70, 4.8)	642 (99, 8.1)
Small	small	503 (53, 5.3)	560 (65, 6.1)	564 (73, 9.5)	626 (93, 3.5)

synthèse

TABLE 7

Neighborhood Density Effects (in Milliseconds) of Number of English (English N) and Dutch Neighbors (English N) for Experiments 1–4 (Effect Size is Calculated by Subtracting RTs in Conditions with a Large Number of Neighbors from Those in Conditions with a Small Number of Neighbors. A Minus Sign Indicates a Facilitatory, and a Plus Sign an Inhibitory Effect.)

		Exp 1	Exp 2	Exp 3	Exp 4	
		Progressive demasking	Progressive demasking	Generalized lexical decision	English lexical decision	
		Blocked	Mixed		Monolinguals	Bilinguals
English	English N	–34	–32	–25	–19	–3
	Dutch N	57	70	27	2	22
Dutch	English N	41	49	8		
	Dutch N	59	70	13		
Nonwords	English N			1	19	19
	Dutch N			26	–5	6

Rôle de la tâche

- On note que l'effet de voisinage interlangue dépend de la tâche
 - *Ex en monolingue: tendance à inhibition en démasquage progressif, facilitation dans certains cas en DL, facilitation importante en naming*
- Démasquage progressif : établissement critère pour identifier 1 mot
 - Joue sur lexique intégré : tous les mots concernés ou pas
- DL: activité lexicale sommée (Grainer & Jacob, 1996) : mots avec bcq de voisins génèrent forte activité lexicale
 - Joue sur lexique intégré : activité sommée pour 1 ou 2 lexiques?
- Pour les nonmots avec bcq voisins, activité lexicale additionnée
 - Joue sur lexique intégré

Modèle BIA

- Modèle algorithmique qui implémente
 - *Traitement bottom up non selectif*
 - Lettres activent mots des 2 langues dans lexique intégré)
 - *Traitement top down langage spécifique*
 - *Avec ces 2 mécanismes, notamment contrôle langue top down, modèle gère résultats sélectifs et non sélectifs*

Par rapport à IA (McClelland & Rumelhart)

- 1 lexique intégré avec mots de 2 langues
- Couche supplémentaire représentationnelle contenant « nœuds de langue »
- Traits activent lettres
- Lettres activent mots (2 langues)
- Relations inhibitrices entre les mots
- Mots activent nœud langue
- Nœud langue activé envoie inhibition aux mots autre langue
- L'activation des nœuds langue reflètent la quantité d'activation dans chaque lexique

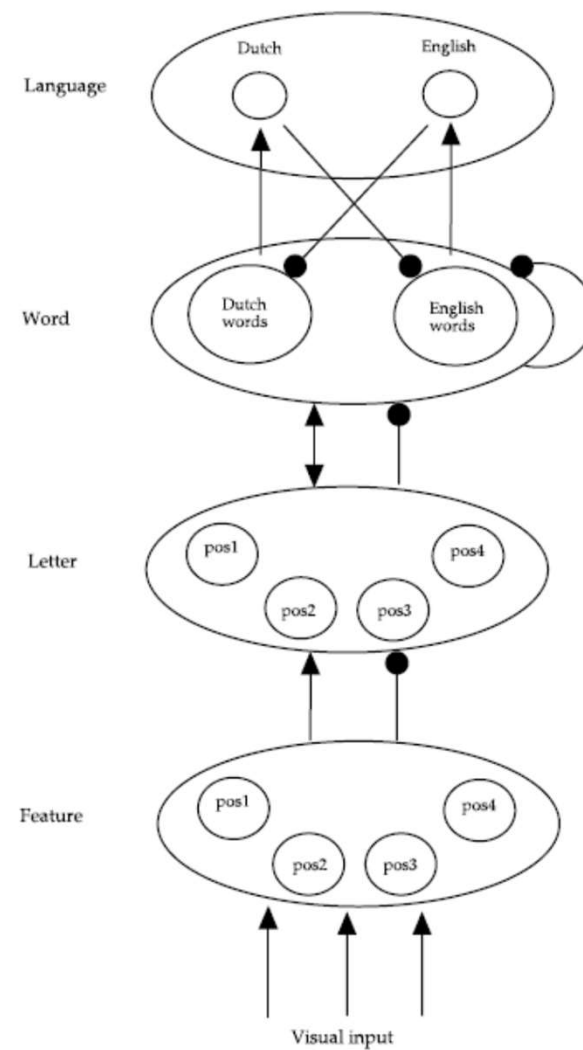


Figure 1. The bilingual interactive activation model. (Adapted from Dijkstra et al., 1998.)

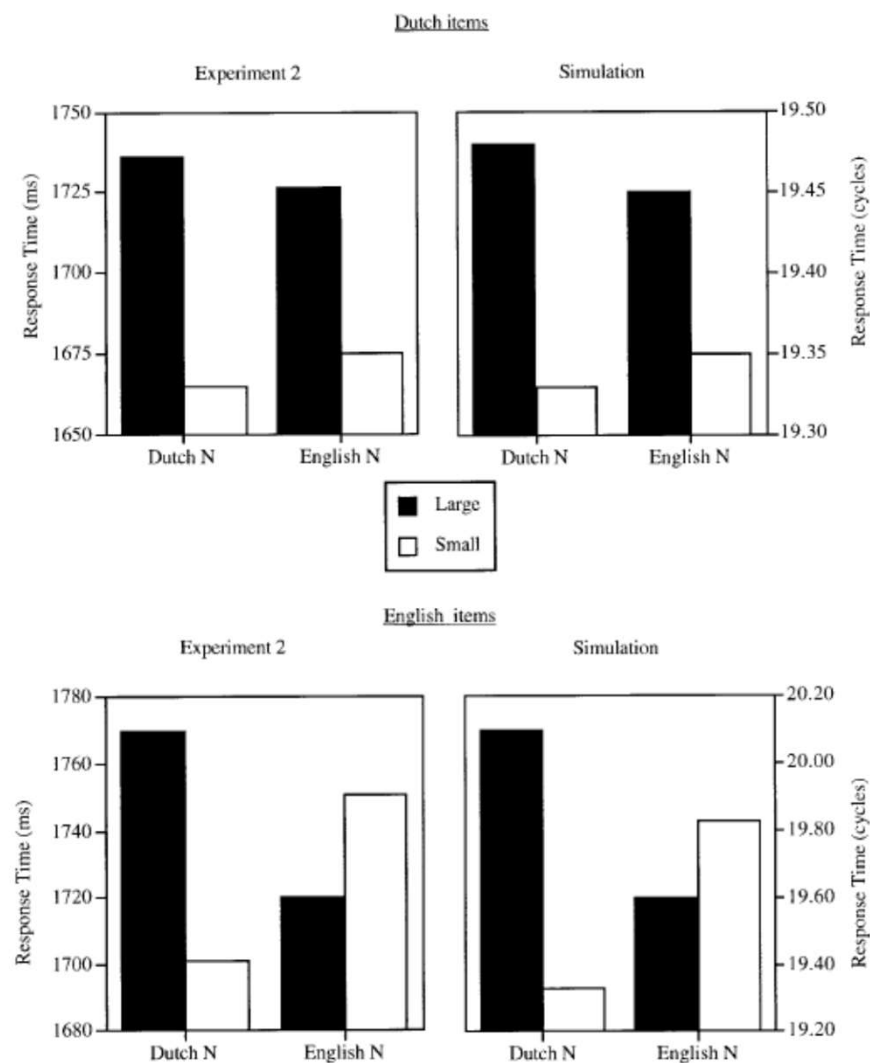


FIG. 3. Mean identification latencies of Experiment 2 (in milliseconds) and simulation results of the BIA model (in cycles) for Dutch N and English N effects on Dutch and English items. In the simulation a reduced upper limit was used for the resting-level activations of English items (-0.30), and the standard IA-model upper limit was used for resting-level activations of Dutch items (0.00). The top-down inhibition parameter from the Dutch language node to all English words was set to 0.03 .

Qq pb posés par BIA/ révision BIA+

- 2 questions critiques
 - *(a) est ce que l'activité des noeuds langue est influence par les sources d'information non linguistiques (en dehors sys RM)*
 - *Est ce qu'à l'intérieur du systeme, les noeuds langue affectent significativement l'activation (et donc la reconnaissance) des mots*
- Le modèle BIA répond OUI aux 2 questions,
- Le modèle BIA+ répond NON aux 2 questions

BIA+

distinction claire et nette entre

- un système d'identification de mots encapsulés et

- un système de tâches/décisions.

Les informations linguistiques provenant du signal d'entrée ou du contexte (de la phrase) peuvent affecter l'identification des mots.

tandis que les informations contextuelles non-linguistiques (par exemple, les attentes et les stratégies des participants) affectent les paramètres du système de tâche/décision.

Le nouveau modèle suppose une interactivité au sein du système d'identification des mots et entre ce système et le système de décision.

Cependant, le modèle propose également que les niveaux d'activation lexicale

au sein du système d'identification des mots lui-même soient

non affectés par le système de tâche/décision

Dans BIA, les nœuds langue

- Servent de tags aux mots (représentations linguistiques)
- Collectent de l'activation
- sont des mécanismes fonctionnels non linguistiques qui régissent l'activation des mots des 2 langues (filtre)
- Y compris potentiellement pré activation contextuelle non linguistique

Pb posés par BIA

- Pas de représentation phonologique ou sémantique
- Représentation homographes interlangues et cognates sous spécifiées
- Confusion entre aspect représentationnel et fonctionnel des noeuds langue
- Peu d'explication de la façon dont context non linguistique et linguistique affecte la reconnaissance
- Pas de lien entre la tâche et la performance

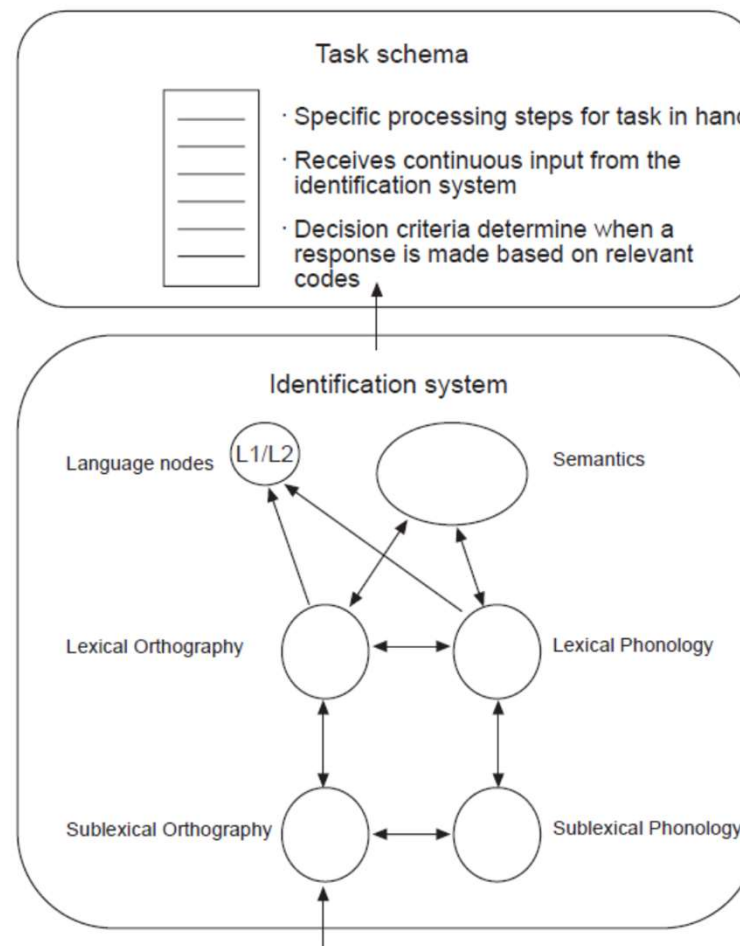


Figure 2. The BIA+ model for bilingual word recognition. Arrows indicate activation flows between representational pools. Inhibitory connections within pools are omitted. Language nodes could instead be attached to lemma representations between word form and meaning representations. Non-linguistic context only affects the task schema level.

Les points modifiés BIA/ BIA+

Représentation
information
phonologique et
sémantique (non
implémentés)

Représentation des
homographes
interlangue

Les nœuds langue

Les effets contextuels
linguistiques

Les effets contextuels
non linguistiques

Les liens entre
reconnaissance de
mots et demandes de
la tâche