

Comment l'information circule d'un niveau de traitement à l'autre lors de l'accès lexical en production verbale de mots ?

Éléments de synthèse

Sébastien Roux¹ et Patrick Bonin^{1,2}

¹LEAD-CNRS (UMR 5022), Université de Bourgogne, Dijon

²Institut universitaire de France

RÉSUMÉ

Cet article propose une synthèse des principaux travaux sur la question de la transmission de l'activation en production verbale à l'oral. Après avoir rappelé brièvement quels sont les différents niveaux des traitement identifiés par les chercheurs, nous décrivons trois types d'architectures fonctionnelles – discrète-sérielle, en cascade et interactive – qui ont été proposées pour rendre compte de la manière dont l'activation se transmet entre les différents niveaux de traitement. Nous exposons ensuite les arguments en faveur de la conception discrète et sérielle soutenue par Levelt, Roelofs et Meyer (1999), tout en les critiquant notamment parce qu'ils reposent le plus souvent sur des absences d'effet. Nous présentons un ensemble de résultats expérimentaux qui appuient fortement une architecture en cascade ou interactive entre niveaux lexical et phonologique en production verbale orale. Nous terminons par une évocation rapide des travaux sur la question des transferts d'activation horizontaux.

The transmission of information from one level to another in speech production

ABSTRACT

In this article, we review the most important studies addressing the issue of the flow of information in speech production. After briefly describing the levels of processing which have been proposed to be involved in speech production, we present three types of cognitive functional architectures that have been put forward in order to account for the activation flow in speech production. We review the evidence which favors the

Correspondance : Patrick Bonin, LEAD-CNRS UMR 5022, Université de Bourgogne, Institut universitaire de France, Pôle AAFE – Esplanade Érasme, BP 26513, 21065 Dijon Cedex. E-mail : patrick.bonin@u-bourgogne.fr. Les auteurs tiennent à remercier Ludovic Ferrand et Nicolas Dumay pour les remarques et critiques constructives émises sur des versions précédentes de cet article.

discrete-serial view of speech production put forward by Levelt, Roelofs and Meyer (1999) while criticising this view since it is often based on null results. We review empirical findings which strongly support the view that information flows in a cascaded or interactive manner between the lexical and phonological levels involved in speech production. Finally, we briefly describe some studies on the issue of information flow above the single word.

INTRODUCTION

Différents modèles de la production du langage existent, surtout à l'oral (par ex. Caramazza, 1997 ; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997 ; Levelt, Roelofs, & Meyer, 1999). Ces modèles ont été élaborés, critiqués, puis raffinés sur la base de données empiriques recueillies à l'aide de différentes méthodes : analyses des erreurs produites par des adultes sains et des patients, mot sur le bout de la langue, expériences de chronométrie mentale chez des adultes (Bonin, 2007 pour une description de ces méthodes). Les modèles de production doivent rendre compte des *mécanismes* et des *représentations* qui sont mobilisés entre l'intention de communiquer une ou plusieurs idées et la réalisation articulatoire (ou graphique) effective. Les modèles diffèrent quant au nombre de niveaux de traitement qu'ils prévoient, à leur caractérisation, mais également et surtout, quant à la façon dont l'information circule d'un niveau à l'autre.

Ces dernières années, de nombreuses recherches ont collecté des données qui favorisent une *conception en cascade* de la circulation de l'information en production verbale, et plus précisément, entre niveaux lexical et phonologique. Nous allons revenir sur ce qu'il faut entendre par lexical et phonologique. Ces données remettent donc fortement en question une conception qui a longtemps été défendue avec force et vigueur essentiellement par Levelt et ses collaborateurs (Levelt, 1989 ; Levelt *et al.*, 1999) : la *conception sérielle et discrète* de la production verbale (voir aussi Butterworth, 1989 ; Garrett, 1975, 1980, 1982). Afin d'écarter toute ambiguïté, il faut d'emblée préciser que le caractère sériel et discret de la transmission de l'information n'a, à l'exception de certaines conceptions princeps de la production verbale orale (par ex. Garrett, 1980), jamais été conçue comme absolu, et ce, même dans la conception discrète et sérielle la plus aboutie, c'est-à-dire celle de Levelt *et al.* (1999). Dans la suite de l'article quand nous parlons de conception discrète et sérielle, nous faisons référence à cette dernière. Dans le modèle WEAVER++ (Roelofs, 2004), qui est une implémentation de la théorie de Levelt *et al.* (1999), il existe des liens bidirectionnels entre les représentations lexicales (les lemmes dans leur acception) et les concepts entre lesquels l'activation se transmet en cascade. Nous soulignons donc que la conception de Levelt *et al.* (1999),

reconnue par tous les psycholinguistiques comme conception discrète et sérielle, prévoit de la transmission en cascade entre certains niveaux de traitement.

Des synthèses en français existent sur la question des niveaux de traitement (Bonin, 2002 ; Ferrand, 1994, 1997), il ne s'agit donc pas, ici, d'en proposer une nouvelle. En revanche, à notre connaissance, il n'existe pas de synthèse récente des recherches centrées sur le problème de la transmission de l'information dans les modèles de production. Ce problème est crucial pour quiconque s'intéresse à la production verbale et il est fort à parier qu'il va continuer à susciter de nombreuses recherches. Il sera surtout question ici de la production verbale orale, non pas que la production verbale écrite ne soit pas importante, mais parce que les travaux sur la question de la circulation de l'information à l'écrit sont encore très peu nombreux. Comme nous le verrons, les arguments en faveur de la conception discrète et sérielle ont principalement été des arguments négatifs. Ainsi, c'est par exemple l'absence d'effet de facilitation ou d'interférence phonologique de voisins sémantiques qui a fourni des arguments en faveur d'une transmission discrète-sérielle de l'information du niveau lemma au niveau des lexèmes. Le plan de l'article est le suivant. Dans une première partie nous allons faire un bref rappel des niveaux de traitement car il n'est pas possible de considérer la circulation de l'information sans préalablement évoquer cette question. Dans une deuxième, nous exposerons les arguments en faveur de la conception discrète et sérielle. Dans une troisième et quatrième partie, nous envisagerons les données de plus en plus nombreuses qui favorisent les conceptions en cascade et interactive respectivement. Nous terminons enfin par une évocation rapide des travaux sur la question des transferts d'activation horizontaux.

I - NIVEAUX DE TRAITEMENT EN PRODUCTION VERBALE ET ARCHITECTURES FONCTIONNELLES

Les modèles visent à rendre compte des mécanismes et des représentations qui sont en jeu entre deux événements qui sont, d'une part l'intention de communiquer une idée ; d'autre part, la production articulatoire (ou graphique). Les chercheurs doivent donc identifier combien il y a de niveau de traitement, les caractériser (la nature des représentations impliquées) et déterminer comment les informations circulent d'un niveau de traitement à l'autre. Autrement dit, les chercheurs ont pour tâche de spécifier une architecture fonctionnelle.

I-1. Des questions autour du nombre et de la nature des niveaux de traitement

Les chercheurs s'accordent sur l'existence de trois niveaux généraux de représentation : conceptuel, lexical et articulatoire (ou graphique). Ils admettent aussi communément la distinction entre un niveau lexical *holistique* (dénommé parfois niveau L, Rapp & Goldrick, 2000) et un niveau phonologique où minimalement les segments individuels sont représentés (cette dernière terminologie est adoptée dans la suite de l'article).

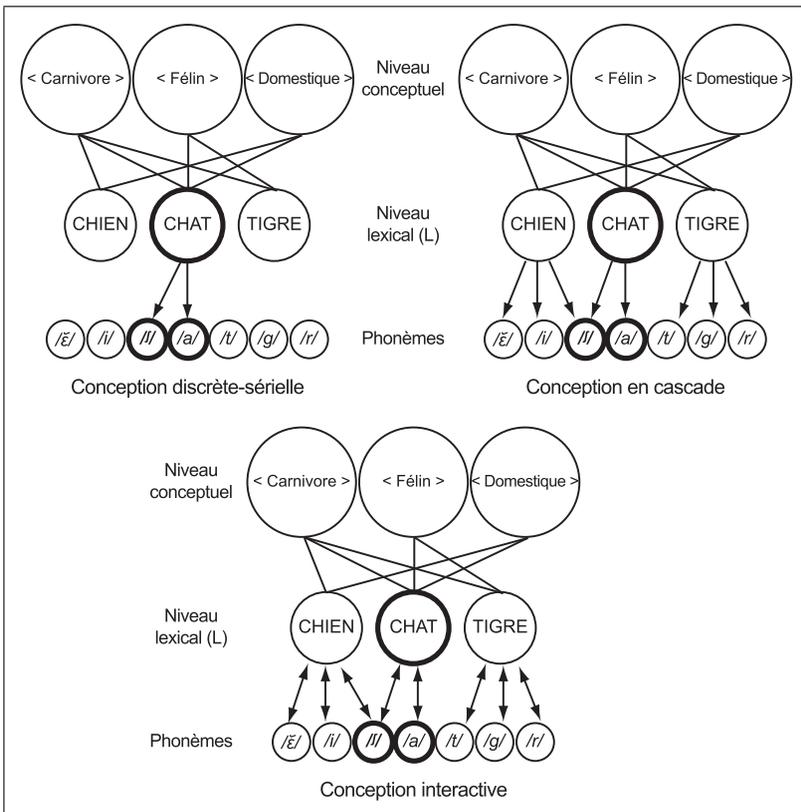


Figure 1. Conception discrète-sérielle (a), en cascade (b) et interactive (c) de la production verbale.

Figure 1. Discrete-serial view (a), cascading view (b) and interactive view (c) of speech production.

L'accord s'arrête toutefois là ! En effet, de nombreuses controverses existent et les problèmes associés ne sont pas résolus. Ainsi, il y a débat sur la nature des représentations sémantiques et sur la façon dont elles contactent les représentations lexicales. Pour illustrer ce problème, prenons le cas de la production orale du mot « père » à partir de la situation d'énonciation d'un protagoniste qui veut signaler à son interlocuteur de se taire car il voit arriver son père dans leur direction. La représentation lexicale correspondant au concept de « PERE » est-elle contactée via l'activation des primitives sémantiques « parent » et « mâle » comme le défendent des chercheurs comme Dell ou Caramazza (Caramazza, 1997 ; Dell, 1986, 1988, 1990 ; Dell & O'Seaghdha, 1991, 1992) ou bien au contraire l'est-elle via un concept unique « PERE » (Roelofs, 2000) ? Un autre problème est celui de savoir si le niveau structural – qui correspond aux connaissances de la forme canonique des objets – fait partie du niveau sémantique/conceptuel (Levelt *et al.*, 1999) ou s'il en est distinct comme le défendent Humphreys et ses collègues (par ex., Humphreys, Riddoch & Quinlan, 1988 ; Humphreys, Lamote & Lloyd-Jones, 1995) ? Doit-on envisager de distinguer une sémantique conceptuelle d'une sémantique lexicale comme le suggère l'observation de patients qui, bien qu'étant incapable de récupérer le nom d'un objet à partir d'une image ou commettent des erreurs sémantiques (par ex., « couteau » à la place de « fourchette ») parviennent néanmoins à se servir de l'objet correctement (Nickels, 2000) ?

Les chercheurs admettent un niveau lexical holistique et un niveau phonologique mais des polémiques ont trait également au nombre et à la nature des niveaux de traitement qu'il faut envisager. Y a-t-il un niveau carrefour obligatoire – le niveau des lemmas (Levelt, 1989 ; Levelt *et al.*, 1999) – entre les représentations sémantiques/conceptuelles et phonologiques (les lexèmes ou les morphèmes selon Levelt *et al.*, 1999) qui spécifie la catégorie grammaticale et le genre pour des langues qui codent le genre ? ou bien ce niveau est-il tout simplement superflu comme l'ont défendu Caramazza (1997) et ses collègues (voir aussi Miozzo et Caramazza, 1997) ? En effet, selon Caramazza (1997), le niveau lexical (L) correspond aux lexèmes phonologiques (et orthographiques en production verbale écrite), lequel donne accès aux segments individuels.

Nous avons présenté ailleurs des données qui permettent d'alimenter la réflexion autour de ces différentes questions (Bonin, 2002). Le problème qui nous intéresse ici est celui de la transmission de l'information d'un niveau de traitement à l'autre : du niveau sémantique/conceptuel au niveau lexical holistique (L), du niveau L au niveau des représentations phonologiques. Comme déjà précisé, il n'existe pas de synthèse récente sur ce problème

en français. Il est bien sûr relié aux questions précédentes mais il peut être traité indépendamment comme nous le faisons dans la suite de l'article.

I-2. Différentes architectures fonctionnelles

Selon Goldrick (2006 ; Rapp et Goldrick, 2000) chaque niveau (ou étape) de traitement débute par la sélection d'une représentation à un niveau N donné (par exemple au niveau sémantique) et se termine par la sélection d'une représentation au niveau (ou étape) subséquente (N +1 ; par exemple au niveau L) dans le flux du traitement. Les niveaux de traitement sont donc balisés par des moments qui correspondent à des *sélections* spécifiques. Par sélection, on entend des mécanismes qui permettent à une ou plusieurs représentations de « dominer » le traitement (Goldrick, 2006). Par exemple, au niveau L, plusieurs représentations lexicales sont activées en parallèle (il s'agit des lemmas selon Levelt *et al.*, 1999) et des mécanismes interviennent à un moment donné pour sélectionner une représentation (le plus souvent celle cible) parmi la cohorte des représentations activées.

Il existe trois types d'architectures fonctionnelles de la production verbale : sérielle-discrète, en cascade et interactive. En théorie, les niveaux de traitement N, N +1, N +2... peuvent être strictement sériels, partiellement sériels (le niveau N démarre le traitement avant N +1 mais se poursuit en parallèle) ou entièrement parallèles (les traitements sont simultanés). Aussi l'autonomie entre les niveaux de traitement peut-elle varier. Lorsqu'un niveau N influence uniquement N +1, et que N +1 influence seulement le niveau subséquent N +2, on parle d'autonomie de traitement ; au contraire, lorsqu'un un niveau de traitement N influence N +2 sans passer par le niveau N +1, on dit que l'autonomie se fait dans ce cas sans médiation.

Dans la littérature sur la production du langage, les chercheurs parlent de transmission sérielle-discrète de l'information lorsque l'activation ne se transmet d'un niveau de traitement N à un niveau N +1 que lorsque le traitement réalisé au niveau N a été entièrement achevé. Plus précisément, selon Goldrick (2006), une architecture sérielle-discrète est définie par trois caractéristiques : 1. le traitement au niveau d'une étape n'est pas initialisé tant que la sélection n'a pas eu lieu à l'étape précédente ; 2. seules les représentations sélectionnées peuvent fournir de l'activation à d'autres étapes de traitement et 3. l'activation circule seulement de « l'amont vers l'aval ». Au contraire, dans une conception en cascade, l'activation circule d'un niveau N à un niveau N +1 en continu de sorte que l'intégralité du traitement N n'est pas achevée lorsque le travail débute au niveau

subséquent $N + 1$. Donc, les représentations à un niveau donné peuvent fournir de l'activation à d'autres niveaux avant que la sélection n'ait eu lieu (Soni, Lambon Ralph, Krist, Ehsan, Hodgson, & Woollams, 2009). Ainsi, le traitement au niveau L peut-il débiter alors même qu'au niveau sémantique le traitement n'est pas achevé.

Une conception interactive est une conception en cascade qui inclut des rétroactions d'un niveau $N + 1$ à un niveau N . Si l'on reprend les critères fournis par Goldrick (2006), une conception interactive est une conception en cascade dans laquelle le traitement s'effectue dans les deux directions, c'est-à-dire de N à $N + 1$ et de $N + 1$ à N . Comme nous allons le présenter, le degré de sérialité versus de cascade est variable dans les différents modèles de production.

Une question qui se pose aussi quant aux architectures fonctionnelles des modèles de production est de savoir si les liens entre les unités qui appartiennent à différents niveaux de traitement, ou au sein d'un même niveau de traitement, peuvent être inhibiteurs ou doivent être considérés comme étant essentiellement excitateurs. La plupart des modèles n'incluent pas de liens inhibiteurs entre unités de traitement et les données empiriques en faveur de l'existence de tels liens sont à l'heure actuelles plutôt ténues.

II - LES TENANTS D'UNE CONCEPTION DISCRÈTE-SÉRIELLE ET LEURS ARGUMENTS

II-1. Où se situe la sérialité dans les modèles de production ?

Comme l'ont étudié Rapp et Goldrick (2000), il est possible d'envisager des modèles où la sérialité est prévue entre tous les niveaux de traitement ; d'autres où la transmission en cascade est restreinte à deux ou trois niveaux de traitement ; d'autres enfin où la transmission (en cascade) se fait au sein de tous les niveaux de traitement.

Plus précisément, Rapp et Goldrick (2000) ont eu recours à des simulations pour tester quatre modèles ayant une architecture fonctionnelle différente de la production verbale : (a) un modèle standard discret ; (b) un modèle en cascade sans interaction ; (c) un modèle interactif restreint et (d) un modèle interactif exhaustif. Ces modèles étaient donc différents selon qu'ils prévoyaient ou non une transmission en cascade

et selon qu'ils admettent ou non des rétroactions entre les niveaux de traitement. Rapp et Goldrick (2000) ont réussi à simuler correctement des données empiriques concernant les erreurs de production (erreurs mixtes « cheval » à la place de « chèvre » ; biais lexical comme « cheval » au lieu du non-mot « chaval » ; erreurs sémantiques « mouton » à la place de « chèvre » et erreurs phonologiques « chaise » pour « chèvre ») uniquement avec un modèle dans lequel il y a une transmission en cascade de l'information et une interactivité seulement entre les niveaux lexical (L) et phonologique. Comme nous l'avons vu, la conception discrète-sérielle prévoit une transmission en cascade entre le niveau des concepts et le niveau des lemmas (niveau L). La transmission est donc discrète et sérielle entre les niveaux L et phonologique.

II-2. Arguments en faveur de la sérialité dans les modèles de production

Au plan théorique, Levelt *et al.* (1991, 1999) ont justifié l'existence de la sérialité dans la production du langage car elle constitue une protection naturelle contre l'émergence d'erreurs pléthoriques. Autrement dit, si la transmission de l'activation du niveau lemma au niveau lexème n'était pas sérielle et discrète, les erreurs seraient bien plus nombreuses chez les locuteurs qu'elles ne le sont. Les statistiques sur les taux d'erreurs chez des adultes montrent qu'elles constituent des événements plutôt rares (moins d'une erreur de sélection lexicale pour 1000 mots produits).

L'observation selon laquelle les échanges de mots (par ex., produire « la porte dans le pied » à la place de « le pied dans la porte ») en production verbale orale chez des adultes sains ont des caractéristiques distinctes des échanges de sons (« rull pouge » produit à la place de « pull rouge ») a été prise comme argument en faveur de la distinction entre les lemmas et les lexèmes et de leur séparation temporelle lors du traitement. Ainsi les échanges de phonèmes (ou segments) concernent-ils en général des unités relevant du même syntagme et de nature grammaticale différente tandis que les erreurs de sélection lexicale impliquent des mots de même catégorie grammaticale qui appartiennent à des syntagmes différents (Garrett, 1975). Il est admis que les échanges de segments se produisent lors de l'encodage phonologique, c'est-à-dire au niveau lexème (ou des morphèmes), alors que les échanges de mots se produisent lors de l'encodage grammatical qui relève du niveau lemma. Toutefois, ce sont les études réalisées chez l'adulte à l'aide des techniques de chronométrie mentale qui ont fourni les arguments les plus forts en faveur de la conception discrète-sérielle.

Une recherche souvent mentionnée dans la littérature en faveur de la conception sérielle-discrète de la production verbale est celle de Schriefers, Meyer et Levelt (1990). Elle a été réalisée à l'aide du paradigme de l'interférence « mot-image ». Dans ce paradigme, une cible imagée est présentée à différents moments (SOAs) avec un distracteur (mot) présenté visuellement ou auditivement. Cette technique permet de recueillir des informations sur la nature des représentations qui sont mobilisées ainsi que sur leur décours temporel d'activation. Dans l'étude de Schriefers *et al.* (1990), des adultes dénommaient oralement des images avec la présentation auditive de distracteurs-mots. Ceux-ci apparaissaient 150 ms avant le début de l'image (SOA = -150 ms), simultanément (SOA = 0 ms), ou 150 ms après (SOA = +150). Les distracteurs étaient reliés aux cibles, soit sémantiquement (« chien-RENARD »), soit phonologiquement (« singe-LINGE »), ou bien ils étaient non-reliés (« pomme-CHAT ») ou encore neutres (la présentation de « blanco »). Les latences de dénomination étaient plus longues avec des distracteurs sémantiquement reliés qu'avec des distracteurs non-reliés tandis qu'elles étaient plus courtes avec des distracteurs phonologiquement reliés qu'avec des distracteurs non-reliés. Toutefois, l'effet d'interférence sémantique était attesté uniquement lorsque les distracteurs étaient présentés à un SOA de -150 ms alors celui de facilitation phonologique apparaissait avec des valeurs de SOA de 0 et de +150 ms. Dans une tâche contrôle de reconnaissance d'images qui, selon Schriefers *et al.* (1990), mobilise le niveau conceptuel mais pas le niveau lexical, l'effet d'interférence sémantique n'était pas observé au SOA de -150 ms. Pour Schriefers *et al.* (1990), l'effet d'interférence sémantique relève du niveau lemma et celui de facilitation phonologique du niveau lexème. Puisque l'effet d'interférence sémantique n'est pas observé dans une tâche de reconnaissance d'objets, par élimination, il ne peut pas être attribué au niveau conceptuel. Puisque ce même effet ne se manifeste pas pour des valeurs de SOAs lors desquelles la facilitation phonologique est observée ; le niveau lexème n'est pas le niveau qui le sous-tend. Les données suggèrent donc que l'accès lexical en production met en jeu les niveaux lemma et lexème et, plus intéressant encore, que ces niveaux ne se chevauchent pas temporellement comme prédit par la conception sérielle-discrète.

Les résultats de Schriefers *et al.* (1990) ne sont toutefois pas aussi clairs que cela quant aux décours temporel d'accès aux représentations lemmas et lexèmes dans la mesure où des études ultérieures ont montré que la facilitation phonologique pouvait être également observée avec l'utilisation de SOA négatifs (Starreveld, 2000). Le paradigme de l'interférence ne fournit pas des données fiables sur le décours temporel de l'activation des

représentations en jeu en production verbale. Enfin, l'interprétation selon laquelle les effets d'interférence sémantique obtenus avec le paradigme de l'interférence traduisent l'activation des lemmas dans la production verbale a été vivement critiquée ces dernières années (par ex., Finkbeiner & Caramazza, 2006 ; Janssen, Schrim, Mahon & Caramazza, 2008 ; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007 ; Miozzo & Caramazza, 2003). Même la facilitation phonologique observée avec le paradigme de l'interférence, et attribuée initialement par Schriefers *et al.* (1990) comme spécifique du niveau lexème, a été réinterprétée par Roelofs, Meyer et Levelt (1996) comme pouvant en partie relever du niveau lemma. Les débats sont donc vifs et dépassent le cadre de cet article mais nous renvoyons le lecteur aux travaux mentionnés pour des détails.

Une autre étude souvent avancée pour soutenir la conception discrète-sérielle est celle de Levelt, Schriefers, Vorberg, Pechmann et Havinga (1991). Dans cette étude, la coactivation phonologique d'items non-cibles a été testée. Dans la conception discrète-sérielle de Levelt *et al.* (1999), l'activation d'un concept cible au niveau préverbal entraîne l'activation en parallèle de plusieurs concepts reliés sémantiquement et des lemmas correspondants (par ex. le concept cible de « CHIEN » active les concepts « CHAT », « RENARD » etc., lesquels activent les lemmas associés). Le traitement au niveau lemma se termine par la sélection d'un seul lemma au sein de la cohorte des lemmas activés. Cette sélection aboutit à ce que les informations lexémiques correspondantes sont récupérées (c'est-à-dire phonèmes individuels, patron intonatif, nombre de syllabes). Autrement dit, la transmission de l'activation du niveau lemma au niveau lexème est sérielle et discrète.

La coactivation phonologique, qui correspond à l'activation simultanée de lexèmes autres que celle du lexème correspondant au lemma sélectionné, n'est donc pas prévue comme le suggèrent les données de Levelt *et al.* (1991). Des participants dénommaient des images (par ex. l'image d'un MOUTON). Pour certains essais-cibles, ils entendaient un item et devaient décider s'il correspondait à un mot ou à un non-mot. Les mots sur lesquels portait la décision lexicale pouvaient être identiques à la cible (« mouton »), sémantiquement reliés (« vache »), phonologiquement reliés (« moulin »), phonologiquement reliés à un voisin sémantique de la cible (« chaise » relié phonologiquement à « chèvre » voisin sémantique de « mouton ») ou non-reliés (« voiture »). Le résultat le plus important était que les mots phonologiquement reliés à un voisin sémantique de la cible (par ex. « chaise ») n'étaient pas traités plus rapidement que ceux non-reliés. Pour Levelt et ses collègues, cette étude montre que les voisins sémantiques d'une cible ne sont pas phonologiquement encodés conformément à ce que prédit une conception discrète-sérielle. Ainsi, c'est l'absence d'effets

phonologiques des voisins sémantiques sur les vitesses de réaction qui permet à Levelt *et al.* (1991) d'argumenter en faveur de la conception sérielle et discrète de la production verbale, ce qui correspond selon nous à un argument « négatif ».

Des chercheurs ont reproché à l'étude de Levelt *et al.* (1991) de s'appuyer sur des mesures (temps de décision lexicale) qui ne sont pas suffisamment fines pour capter des effets de coactivation phonologiques, effets qui sont en principe de faible ampleur (Dell, & O'Seaghdha, 1991). Cependant, une étude de Jescheniak, Hahne et Schriefers (2003) s'appuyant sur des mesures fines, c'est-à-dire des potentiels évoqués, a permis d'obtenir un patron de résultats conforme à celui de Levelt *et al.* (1991) en regard de la coactivation phonologique. Néanmoins, en ayant recours au paradigme de l'interférence avec plusieurs distracteurs reliés phonologiquement aux voisins sémantiques de cibles à produire en dénomination, Abdel Rahman et Melinger (2008) ont réussi à mettre en évidence sur les latences de dénomination des effets de coactivation phonologique. Par exemple, l'image d'un MOUTON était présentée avec les deux distracteurs-mots « chine » et « lèvre ». Ces mots distracteurs sont reliés phonologiquement à « chèvre », qui est un voisin sémantique de mouton. Conformément à ce que prédit une conception en cascade, des effets inhibiteurs étaient observés. Il est donc clair que l'absence d'effets de coactivation phonologiques mise en avant par Levelt *et al.* (1991) pour réfuter une conception en cascade est un argument négatif qui ne vaut que tant que l'absence d'effet en question n'a pas été levée par le recueil de données empiriques pertinentes.

La théorie de Levelt et collaborateurs (1999) a toutefois été remaniée quant au principe de transmission sériel et discret de l'activation du niveau lemma au niveau lexème (Jescheniak & Schriefers, 1999). La coactivation phonologique se produit lorsque deux items lexicaux sont contextuellement appropriés pour exprimer un concept donné. Pour Levelt, lexicalement approprié signifie que le terme lexical transmet de façon pertinente l'idée cible. Ainsi, le terme « auto » est-il autant pertinent que son synonyme « voiture » pour transmettre l'idée correspondante. Pourquoi un tel assouplissement ? Il a été introduit pour rendre compte de données expérimentales qui montrent que des quasi-synonymes (par ex. « voiture » et « automobile ») font l'objet d'une coactivation phonologique (Jescheniak & Schriefers, 1997, 1998 ; Peterson & Savoy, 1998). La transmission de l'information serait donc par défaut discrète et sérielle mais le système tolérerait quelques exceptions comme les synonymes.

Malgré cet ensemble de résultats en faveur de la conception discrète-sérielle, qui comme nous l'avons souligné, sont pour l'essentiel des résultats négatifs, il existe d'autres études, nombreuses, qui s'accordent mieux avec

une conception en cascade de la production verbale. Nous allons nous attacher à présenter certaines d'entre elles dans la partie suivante.

III - CONCEPTIONS EN CASCADE ET INTERACTIVE : ARGUMENTS

Parmi les modèles qui prévoient une transmission en cascade, il y a le modèle d'Humphreys et collègues (Humphreys *et al.*, 1988) qui n' a cessé d'être testé pendant de nombreuses années et qui est sans doute l'un des premiers modèles en cascade de la production des plus explicites. Nous allons brièvement présenter ce modèle afin de permettre au lecteur de mieux comprendre les arguments qui le fondent.

Ce modèle comporte trois niveaux principaux de traitement – structural, sémantique et phonologique – entre lesquels l'activation est transmise de façon continue (pour une implémentation de ce modèle voir Humphreys, Lamote, & Lloyd-Jones, 1995). Des liens excitateurs et inhibiteurs existent entre unités appartenant à des niveaux de représentation distincts. Aussi des liens inhibiteurs sont-ils prévus au sein des unités elles-mêmes qui relèvent des niveaux structural et phonologique respectivement. La dynamique fonctionnelle fait que le système de production atteint un état stable lorsque les représentations qui correspondent à l'objet à dénommer sont fortement activées tandis que celles qui sont concurrentes sont inhibées. Ce modèle ne prévoit pas de rétroactions entre niveaux de traitement : il est donc strictement « en cascade ».

D'autres modèles en cascade ont été proposés comme celui de Caramazza (1997). L'intérêt de ce modèle est qu'il n'est pas restreint à la modalité de production à l'oral. Par ailleurs, comme déjà précisé, ce modèle exclut le niveau des lemmas : les informations sémantiques/conceptuelles sont liées directement aux représentations lexicales phonologiques. Les informations sur le genre et la catégorie grammaticale sont codées au niveau phonologique (lexèmes).

III - 1. De quelques interactions en faveur de la cascade

Le modèle de Humphreys *et al.* (1988) prédit qu'une compétition au niveau structural entre des unités co-activées va se répercuter sur les niveaux subséquents, et donc au niveau lexical puisque l'information se transmet en continu. Au contraire, dans une conception discrète, des effets qui se

produisent au niveau de l'accès aux représentations structurales doivent s'additionner avec ceux qui affectent un niveau de traitement subséquent comme celui lexical (*cf.* logique de Sternberg, 1969). Ainsi, les effets de la fréquence qui sont généralement interprétés comme relevant du niveau lexical (par ex. Jescheniak & Levelt, 1994 mais voir Knobel, Finkbeiner & Caramazza, 2008) ne doivent pas varier pour des objets appartenant à des catégories « perceptivement » similaires versus dissimilaires dans la mesure où l'accès au nom des objets ne s'effectue que lorsque les traitements réalisés au niveau structural sont achevés. Humphreys *et al.* (1988) ont mis en évidence dans une tâche de dénomination que l'effet de fréquence (les items rares sont traités plus rapidement que ceux fréquents) était plus important pour des items de catégories dont les membres sont « perceptivement » dissimilaires (par ex. les outils) que pour des items de catégories perceptivement similaires (par ex. les fruits). Ces données s'accordent avec une conception en cascade dans laquelle les représentations lexicales sont activées avant même que l'accès et la sélection de représentations structurales ne soient terminés. Il est intéressant de remarquer que ce résultat est clairement embarrassant pour Levelt *et al.* (1999) comme ils l'ont d'ailleurs reconnu : *Clearly, our serial stage model does not predict this interaction of a conceptual variable with name frequency* (p. 37) (Clairement, notre modèle en étape sérielle ne prédit pas cette interaction entre une variable conceptuelle et la fréquence lexicale, notre traduction). Mais Levelt *et al.* (1999) se sont défendus en mettant en avant une possible variable confondue. La fréquence lexicale pourrait être confondue avec la familiarité conceptuelle qui relève donc du niveau conceptuel. Si la fréquence lexicale est véritablement confondue avec la familiarité conceptuelle, l'interaction rapportée par Humphreys *et al.* (1988) n'est plus critique pour remettre en question la conception discrète-sérielle de Levelt. Ce type d'argument est encore une fois un argument négatif ! De plus, pour que la critique soit effective, Levelt *et al.* (1999) se devaient de montrer que la fréquence lexicale était effectivement confondue avec la familiarité conceptuelle dans l'étude de Humphreys *et al.* (1988).

Vitkovitch et Humphreys (1991) ont fait dénommer des adultes en les soumettant à une pression temporelle. Pour réaliser cela, ils ont eu recours à une technique qui nécessite des participants qu'ils dénomment des images avant l'émission d'un signal sonore qui apparaît peu de temps après la présentation d'une cible imagée. Dans cette situation, les participants ont produit un nombre d'erreurs beaucoup plus élevé que dans une situation sans pression temporelle. Un type d'erreurs relevé par Vitkovitch et Humphreys (1991) consistait en des persévérations : les individus produisaient le nom d'une image qui avait été préalablement dénommée (par ex. le mot « chat » à la

présentation du dessin d'une SOURIS). De plus, Vitkovitch, Humphreys et Lloyd-Jones (1993) ont montré que les erreurs de persévération observées dans des conditions de dénomination sous pression temporelle étaient toujours reliées sémantiquement et visuellement aux images cibles. Cette observation s'accorde avec l'hypothèse selon laquelle le recouvrement du nom des images est contraint par la similarité visuelle et sémantique entre objets, et donc, que l'activation se transmet en cascade.

Toujours dans la même veine, Griffin et Bock (1998) ont montré que les latences de dénomination étaient plus courtes pour des images ayant un label fréquent que pour celles ayant un label rare lorsque la dénomination s'effectuait dans un contexte où les participants ne pouvaient pas prédire l'apparition des images (par ex., avant la présentation de l'image d'un TRAIN, on présente la phrase contexte « Vous allez voir apparaître sur l'écran l'image d'un... »). En revanche, cet effet de fréquence s'inversait de manière non significative lorsque la dénomination se réalisait dans un contexte prédictif (par ex., la phrase contexte est par exemple « En retard, il se dirigeait à vive allure vers la gare pour prendre son... »). Nous avons observé un tel effet en production verbale écrite (Bonin, Roux, Barry, & Larell, en préparation), ce qui suggère donc, pour les deux modalités de production, une transmission en cascade de l'information du niveau sémantique au niveau lexical.

III - 2. L'activation phonologique de concepts non-cibles

Un aspect essentiel qui oppose les conceptions discrète-sérielle et en cascade de la production est l'activation phonologique de concepts non sélectionnés. Dans une conception sérielle et discrète, comme celle de Levelt *et al.* (1999), l'activation de la phonologie est restreinte à la représentation conceptuelle sélectionnée, tandis que dans une conception en cascade comme celle de Humphreys *et al.* (1988), l'activation phonologique n'est pas limitée à la représentation conceptuelle sélectionnée : elle se produit également pour des items qui se trouvent dans la situation d'énonciation mais qui ne sont pas sélectionnés. Pour illustrer, imaginons que nous souhaitions communiquer oralement le concept de « cendrier », dont l'objet correspondant se trouve physiquement sur une table. À côté du cendrier, il y a un livre. Dans une conception en cascade, même si le parleur « choisit » de ne produire verbalement que « cendrier » au sein d'une expression telle que « donne moi le cendrier s'il te plaît », dès lors que l'item « livre » se trouve être dans la situation communicative, de par sa contiguïté physique avec le cendrier, le concept correspondant est activé ainsi qu'automatiquement sa représentation phonologique.

Cet aspect a été récemment étudié chez des adultes au moyen du paradigme de la dénomination d'images superposées, ainsi que chez des adultes bilingues dans des tâches de dénomination d'images « cognates ». Les cognates correspondent à des noms qui sont phonologiquement et orthographiquement reliés d'une langue à l'autre comme *gato* qui veut dire « chat » en espagnol et son équivalent catalan *gat* ou encore « lampe » et *lamp* en français et en anglais respectivement. Nous allons maintenant décrire les principales études réalisées récemment à l'aide de ces deux paradigmes.

Dénomination d'images superposées

Dans ce paradigme, des paires d'images superposées sont présentées aux participants (voir Figure 2 pour une illustration). Au sein de chaque paire, un dessin en traits verts fait office de cible et un autre en traits rouges a le statut de distracteur. Le participant doit dénommer le plus rapidement possible l'image « verte », tout en ignorant l'image « rouge ». Ce paradigme correspond donc à une version image (cible)-image (distracteur) du paradigme de l'interférence. L'image distractive peut être soit phonologiquement reliée à l'image cible (par exemple le couple « CANARD-camion »), soit non reliée (« CANARD-pomme ») comme l'illustre la Figure 2.

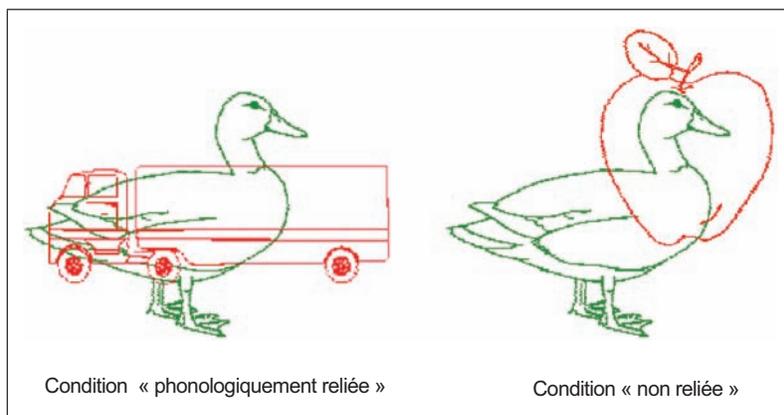


Figure 2. Exemple d'une paire d'images dont les labels sont phonologiquement reliés et d'une autre paire dont les labels sont non-reliés.

Figure 2. Examples of a picture pair whose names are phonologically related and of a picture pair whose names are unrelated.

Ce paradigme a permis de tester des prédictions qui dérivent d'une conception en cascade. Notamment, une prédiction est que si l'activation se transmet en cascade du niveau conceptuel au niveau des représentations phonologiques, les phonèmes d'un item cible (par ex. « canard ») doivent être activés plus rapidement quand ils reçoivent une activation supplémentaire en provenance de l'item distracteur que lorsque la situation ne le permet pas (Figure 3).

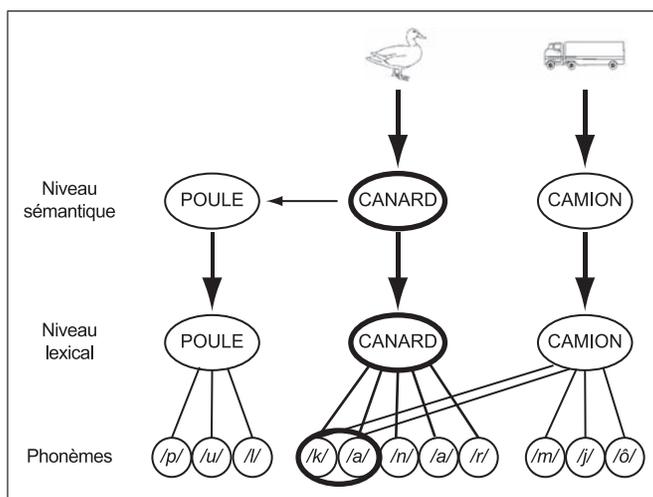


Figure 3. Illustration de la prédiction dérivée d'une conception en cascade et testée avec le paradigme de l'interférence image-image.

Figure 3. Prediction from a cascading view of speech production tested with the picture-picture interference paradigm.

Par exemple, l'item cible CANARD active les phonèmes /k/, /a/, /n/ etc. Si l'item distracteur phonologiquement relié (« camion ») active lui aussi sa phonologie, les phonèmes /k/ et /a/ reçoivent de l'activation en provenance des représentations lexicales « canard » et « camion ». Au contraire, avec une distracteur non-relié comme « pomme », les phonèmes /k/ et /a/ du mot cible ne bénéficient pas d'un supplément d'activation puisqu'ils ne sont pas partagés avec ceux de l'item distracteur. La phonologie de l'image cible doit donc être activée plus rapidement en condition phonologiquement reliée qu'en condition non-reliée. Dans le cadre une conception sérielle et discrète entre niveaux lexical et phonologique (Levelt *et al.*, 1999), puisque seule la représentation lexicale sélectionnée contacte sa phonologie, le fait que celle-ci soit partagée par le label de l'image

distractrice ne doit pas influencer la dénomination de l'image cible. En effet, dans la conception de Levelt *et al.* (1999), le concept sélectionné au niveau préverbal active son lemma correspondant. Toutefois, d'autres lemmas peuvent être activés en parallèle du fait de l'existence de liens au niveau conceptuel. Cependant, seul le lemma correspondant à l'intention de communication est sélectionné. À partir de là, la seule phonologie active est celle de son lemma correspondant. En aucun cas, la phonologie des lemmas qui ont été préalablement coactivés n'entre en jeu.

Morsella et Miozzo (2002) ont utilisé le paradigme de superposition d'images chez des participants anglophones. Ils devaient dénommer le plus rapidement et le plus précisément possible une image cible tout en s'efforçant d'ignorer une image distractrice. Les labels des couples d'images étaient soit reliés phonologiquement au niveau de l'attaque (par ex. « *BED-bell* »), soit non reliés (« *BED-cat* »). Les latences de dénomination orales étaient plus courtes dans la condition phonologiquement reliée que dans celle non-reliée conformément à une conception en cascade. Afin d'écartier une interprétation selon laquelle l'effet serait imputable à des difficultés de traitement perceptif des stimuli (les couples reliés seraient visuellement moins complexes que ceux non-reliés), les mêmes paires d'images ont été dénommées par des locuteurs italiens dans une autre expérience. Ainsi, les paires qui étaient phonologiquement reliées pour des locuteurs anglais (« *BED-bell* ») s'avéraient non-reliées pour des locuteurs italiens (« *BED-bell* » devient « *LETTO-campana* »). Si la différence de latences observées chez les locuteurs anglais était due à une identification perceptive plus rapide des paires d'images reliées phonologiquement par rapport à celles non-reliées, cet effet devrait se manifester indépendamment de la langue de production. Cependant les latences de dénomination des locuteurs italiens se sont révélées quasi-identiques dans les deux conditions expérimentales. Pour Morsella et Miozzo (2002), l'effet de facilitation phonologique signe donc l'activation phonologique de l'item représenté par l'image distractrice. Cette étude fournit un argument fort en faveur d'une transmission en cascade entre niveaux lexical et phonologique. L'effet de facilitation phonologique a été répliqué en anglais par Meyer et Damian (2007). De plus, un effet facilitateur de même taille a été obtenu avec un recouvrement phonologique de la rime (par ex. « *HORN-corn* » / « *CORNE-grain* ») comparativement à un recouvrement de l'attaque (« *CLOUD-clown* » / « *NUAGE-clown* »).

D'autres résultats obtenus grâce au paradigme de l'interférence image-image nuancent toutefois ceux de Morsella et Miozzo (2002). Dans une expérience conduite en allemand, Damian et Bowers (2003) présentaient des paires d'images (superposées) soit sémantiquement reliées

(par ex. « LIT-chaise »), soit non-relies (« LIT-robe »). Les participants devaient à chaque fois dénommer l'une des deux et ignorer l'autre. De façon surprenante, dans cette situation, les latences de dénomination ne différaient pas significativement entre les conditions « reliée » et « non-reliee » alors que l'effet d'interférence sémantique était attesté lorsque les images distractrices étaient remplacées par les mots correspondants (paradigme de l'interférence image-mot). Pour Damian et Bowers (2003), l'effet d'interférence sémantique avec des distracteurs mots serait dû à une compétition au niveau lexical (mais voir Mahon *et al.*, 2007). L'absence d'effet d'interférence sémantique avec des distracteurs-images suggère que la représentation lexicale de l'image distractrice n'est pas récupérée et n'entre donc pas en compétition avec celle de l'image cible. Si l'on admet cette explication, l'hypothèse d'une activation phonologique « automatique » des items non-cibles défendue par Morsella et Miozzo (2002) ne tient plus.

Navarette et Costa (2005) ont proposé une interprétation qui réconcilie l'absence d'effet d'interférence sémantique et la présence d'une activation phonologique dans le paradigme de l'interférence image-image. Pour eux, l'absence d'une interférence sémantique lorsque le distracteur est une image est due à l'action combinée de deux effets antagonistes : une facilitation conceptuelle et une interférence lexicale. Une relation sémantique entre images cible (« tigre ») et distractrice (« lion ») facilite la récupération de la représentation conceptuelle de l'image cible. La représentation conceptuelle de la cible et celle du distracteur transmettent ensuite de l'activation à leurs représentations lexicales respectives. Un concept sémantiquement relié à la cible (par ex. LION) étant plus activé qu'un concept non sémantiquement relié (par ex. PORTE), il transmet davantage d'activation à sa représentation lexicale ; cette dernière entre donc plus fortement en compétition que celle associée à l'image non-reliee (PORTE). La sélection lexicale est donc plus difficile en présence d'un compétiteur sémantiquement relié qu'en présence d'un compétiteur non-reliee. Au niveau lexical, un lien sémantique entre cible et distracteur exerce un effet inhibiteur. Ainsi, dans le cadre du paradigme de l'interférence image-image, l'absence d'effet significatif de l'interférence sémantique dans l'étude de Damian et Bowers (2003) n'est pas incompatible avec la facilitation phonologique observée par Morsella et Miozzo (2002).

L'hypothèse de l'activation phonologique *automatique* de l'image contexte a été récemment remise en question par Jescheniak, Oppermann, Hantsch, Wagner et Mädebach (2009). Ces chercheurs ont en effet échoué à reproduire en allemand les résultats de Morsella et Miozzo (2002). Toutefois, leurs expériences diffèrent légèrement de cette étude princeps.

En effet, les paires d'images restaient affichées à l'écran pendant seulement 1 000 ms (ou jusqu'à ce que le participant donne sa réponse) alors que dans les études de Morsella et Miozzo (2002) et Meyer et Damian (2007), les participants disposaient respectivement de 2 000 ms et 3 000 ms¹. On peut conjecturer que les participants de l'étude de Jescheniak *et al.* (2008) ont ressenti une relative pression temporelle pour dénommer et s'y sont adaptés. Il a pu en résulter que l'image contexte n'était pas suffisamment traitée. Une telle interprétation permet de réconcilier les résultats de Jescheniak *et al.* (2008) avec ceux qui montrent une activation phonologique des images contextes. Elle reste toutefois à tester d'une façon plus systématique.

Récemment, Oppermann, Jescheniak et Schriefers (2008) ont suggéré que l'activation phonologique de l'image contexte n'a lieu que lorsque les images cible et contexte sont imbriquées dans une scène conceptuellement cohérente. Ils présentaient à des locuteurs allemands des couples d'images juxtaposées formant soit une scène conceptuellement cohérente (par ex. *Die Maus frisst den Käse* = « la souris mange le fromage » ; Expériences 1, 2 et 3) soit une scène conceptuellement incohérente (par ex. « le fromage est à côté d'un doigt »). Les participants devaient dénommer l'une des deux images en fonction soit de son rôle thématique (agent *versus* patient dans l'Expérience 1), soit de sa couleur (Expériences 2, 3 et 4). En même temps que les images, étaient présentés auditivement des distracteurs-mots qui étaient soit phonologiquement reliés au label de l'image cible, soit phonologiquement reliés au label de l'image contexte, soit non-reliés. Par exemple, pour la scène représentant une souris mangeant du fromage, le distracteur pouvait être phonologiquement relié à la cible SOURIS (« Mauk » pour *Maus* en allemand), soit phonologiquement relié au distracteur « fromage » (« Kära » pour *Käse* en allemand), soit un distracteur relié à aucun des deux labels (« Pafni »). Lorsque les images cible et contexte constituaient une scène conceptuellement cohérente, Oppermann *et al.* (2008) ont observé des latences de dénomination plus longues avec un distracteur phonologiquement relié à l'image contexte qu'avec un distracteur non-relié. Cet effet d'interférence signe l'activation phonologique de l'image contexte. Cependant, lorsque les deux images ne formaient pas une scène cohérente, il n'y avait plus de différence

¹On peut cependant objecter que Costa *et al.* (2005) ont obtenu des effets de facilitation phonologiques malgré une présentation des paires d'images pendant une durée maximum de 800 ms (c'est-à-dire inférieure à celle de Jescheniak *et al.*, 2008). Toutefois, la latence de dénomination moyenne des participants de Costa *et al.* (2005 ; Expérience 1) est supérieure à celle des participants de Jescheniak *et al.* (2008). Cela suggère que les participants ont été moins sensibles à la pression temporelle dans l'Expérience 1 de Costa *et al.* (2005) que dans celles de Jescheniak *et al.* (2008).

significative entre distracteurs phonologiquement reliés à l'image contexte et non-reliés. Pour ces chercheurs, l'activation phonologique de l'image contexte n'a lieu que lorsque cette dernière forme avec l'image cible une scène conceptuellement cohérente. Les résultats de Jescheniak *et al.* (2008) semblent donc difficiles à réconcilier avec l'hypothèse d'une activation automatique de la phonologie de l'image contexte défendue par Morsella et Miozzo (2002 ; Navarrete & Costa, 2005 ; Meyer & Damian, 2007). Toutefois, Oppermann et ses collègues ont nuancé leur interprétation en suggérant que leur paradigme, différent de celui utilisé par Morsella et Miozzo (2002), n'était peut-être pas assez sensible, et révélait l'activation phonologique de l'image contexte uniquement dans les situations où elle était accentuée par la congruence conceptuelle.

Les résultats qui montrent une activation phonologique d'items non-cibles en production verbale ne sont également pas compatibles avec la conception de Bloem et LaHeij (2003) selon laquelle la production se traduit par la sélection d'un concept cible au niveau du message préverbal. Ils ont défendu cette position sur la base des résultats suivants. Des adultes devaient traduire oralement des mots anglais en allemand, tout en ignorant la présentation d'un distracteur comme l'illustre la Figure 4.

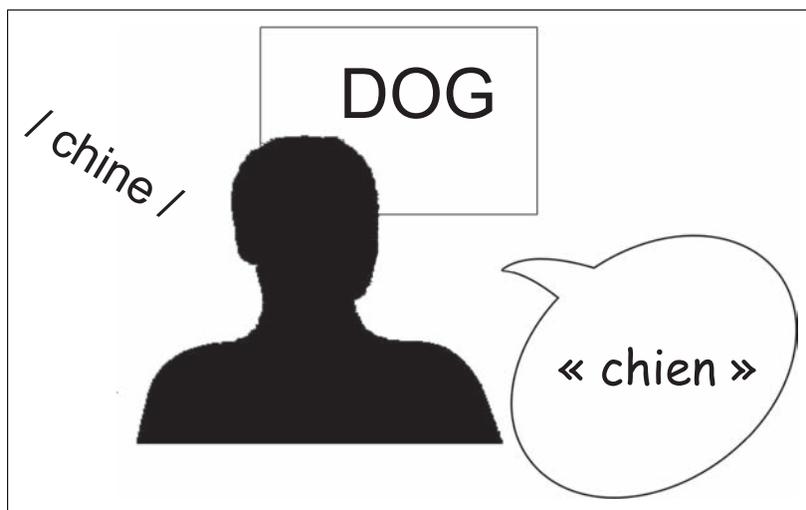


Figure 4. Illustration de la tâche de traduction avec distracteur auditif utilisée par Bloem et LaHeij (2003).

Figure 4. Translation task with an auditory distractor used by Bloem and LaHeij (2003).

Le distracteur pouvait être soit un mot écrit en allemand soit une image. De plus, il était ou non phonologiquement relié à la traduction en allemand du mot cible. Comme l'illustre la Figure 4 (avec un exemple adapté au français où le mot *dog*, donc « chien » en français, est accompagné du distracteur-mot « chine » phonologiquement relié à « chien »), les participants devaient traduire le mot cible (par ex. *arrow* = « flèche ») en allemand (*Pijl*) tout en s'efforçant d'ignorer soit le distracteur-mot (ou image) phonologiquement relié *Pijp* (« pipe »), soit le distracteur-mot (ou image) non-relié *Lepel* (« cuillère »). Avec des distracteurs-mots, les cibles étaient traduites plus rapidement en condition phonologiquement reliée qu'en condition non-reliée. Aucune différence n'était observée avec des distracteurs correspondant à des images. L'absence de facilitation phonologique lorsque le distracteur est une image pourrait être due à ce que la représentation phonologique prend plus de temps à s'activer à partir d'un stimulus imagé qu'à partir d'un stimulus-mot. Selon cette hypothèse, un effet de facilitation phonologique pourrait donc survenir si l'image distractrice phonologiquement reliée était présentée plus longuement pour permettre l'activation de sa représentation lexicale. Cependant, même avec un SOA de -250 ms (c'est-à-dire l'image distractrice précède le mot cible de 250 ms), les mots cibles n'étaient pas traduits plus rapidement dans la condition phonologiquement reliée que dans celle non-reliée (ce délai est normalement suffisant pour permettre une activation phonologique). Ces résultats ne s'accordent donc pas avec l'hypothèse d'une activation automatique de la phonologie d'une image non-cible. Navarette et Costa (2005) ont toutefois proposé que l'absence d'effet phonologique serait due au coût attentionnel élevé de la tâche de traduction. Ce coût ne permettrait pas de traiter l'image distractrice « profondément », de sorte que l'activation générée à partir de l'image serait restreinte au niveau conceptuel. Donc, elle ne se transmettrait pas aux niveaux de traitement ultérieurs. Cette interprétation permet de rendre compte à la fois de l'effet de facilitation phonologique obtenu par Morsella et Miozzo (2002) et de l'absence d'effet rapporté par Bloem et LaHeij (2003).

D'autres résultats corroborent l'hypothèse d'activation automatique de la phonologie d'items non-cibles. Navarette et Costa (2005) ont répliqué l'effet de facilitation phonologique observé par Morsella et Miozzo (2002) en espagnol avec le paradigme de l'interférence image-image. Ils l'ont également obtenu dans une tâche de dénomination de couleur : les images étaient en tracés de couleur rouge, orange, marron ou vert et les participants dénommaient la couleur des traits. Dans la condition reliée, le label de l'image et le nom de la couleur partageaient au moins les deux premiers phonèmes (par ex. « *VELA/verde* » = une BOUGIE de couleur « verte »)

tandis que dans celle non-reliée, le label de l'image et le nom de la couleur ne partageaient aucun phonème (par ex. « *VELA/rojo* » = une BOUGIE de couleur « rouge »). Les latences étaient plus courtes lorsque la couleur était reliée au label de l'image que lorsqu'elle ne l'était pas. Ce résultat suggère que les représentations lexicales du nom de la couleur et de l'image ont transmis de l'activation à leurs phonèmes respectifs.

En résumé, l'effet de facilitation phonologique a été mis en évidence dans plusieurs études, et ce, principalement à l'aide du paradigme de l'interférence image-image, ce qui suggère fortement que des items non-cibles activent automatiquement leur phonologie, contrairement à ce que prévoit une conception sérielle et discrète de l'accès lexical en production verbale. Ces résultats récents favorisent donc clairement une conception en cascade.

Dénomination d'images « cognates »

Une seconde série d'études favorise une conception en cascade de la production verbale. Il s'agit d'études conduites chez des adultes bilingues. Pour appréhender la question de l'activation automatique de la phonologie d'items non-cibles, les chercheurs font dénommer des images dont les noms dans une langue (L1) sont phonologiquement reliés à une autre (L2). Par exemple, un chat en espagnol se dit *gato* et ce nom est phonologiquement relié à son équivalent catalan *gat*. Les images dont les noms dans une langue L1 sont phonologiquement reliés à une autre L2 sont désignées dans la suite de l'article « images cognates ». L'observation d'une facilitation en production d'images cognates s'accorde avec l'hypothèse d'une transmission automatique de la phonologie en production verbale d'items non-cibles (ceux qui relèvent de la langue non sélectionnée chez des locuteurs). Ces données favorisent une conception en cascade de la production.

Costa, Caramazza et Sebastian-Gallès (2000) ont manipulé le statut des images dans une tâche de dénomination orale d'images en espagnol chez des locuteurs bilingues espagnol-catalan et chez des monolingues espagnols : les images pouvaient être « cognates » ou au contraire non-cognates. Les images cognates étaient dénommées plus rapidement que les images-contrôles (non-cognates) par les bilingues uniquement. L'absence de différence significative entre les deux types de stimuli chez les monolingues permet de rejeter une interprétation de l'effet de facilitation phonologique en termes de complexité visuelle. L'effet de facilitation lié au statut cognate des images a été attesté dans d'autres études avec des bilingues de différentes langues (Janssen, 1999 ; Kroll, Dijkstra, Janssen &

Schriefers, 1999, bilingues allemand-anglais et anglais-français ; Gollan & Acenas, 2004, bilingues espagnol-anglais et philipin-anglais ; Hoshino & Kroll, 2008, bilingues espagnol-anglais et japonais-anglais). Ces résultats ont été étendus à des mesures électrophysiologiques (Christofels, Frik, & Schiller, 2007).

La question de l'origine des effets de facilitation liés au statut cognate des items en dénomination d'images est largement débattue. Van Hell et De Groot (1998) ont proposé un *locus* conceptuel de ces effets. Pour eux, les images cognates bénéficieraient d'un plus large recouvrement conceptuel que celles non-cognates, de sorte que l'accès aux représentations sémantiques des images cognates serait plus rapide que celui des images non-cognates. Cette interprétation rend compte des différences de performance entre ces deux types d'images dans des tâches de dénomination orale. Toutefois, d'autres résultats expérimentaux résistent à cette interprétation. Gollan et Acenas (2004) ont utilisé le statut cognate d'images à dénommer pour étudier le phénomène du mot sur le bout de la langue (MBL). Le nombre de MBL est plus faible avec des images cognates qu'avec des non-cognates chez des bilingues espagnol-anglais, et ce, alors même qu'en état de MBL ils peuvent accéder aux représentations sémantiques qui correspondent aux deux types d'images. Le plus fort taux de MBL pour les images non-cognates ne serait donc pas lié à une difficulté d'accès aux représentations sémantiques. Les résultats de Costa *et al.* (2000 ; Expérience 2) posent également problème à un *locus* conceptuel de l'effet facilitateur du statut cognate en dénomination orale. En effet, ils ont observé que la langue de production modulait l'effet du statut cognate dans une tâche de dénomination orale d'images chez des bilingues espagnol-catalan. En fait, la taille de l'effet de facilitation du statut cognate était plus important lorsque les bilingues dénommaient les images dans leur langue non-dominante (le catalan dans leur étude) que dans leur langue dominante (l'espagnol). Or, si l'effet du statut cognate des images était lié à l'accès aux représentations sémantiques, il devrait être identique dans les deux langues.

Costa *et al.* (2000) proposent que l'effet facilitateur lié aux cognates est d'origine lexicale, de sorte qu'il s'accorde avec une architecture en cascade de la production verbale orale chez des bilingues. Dans ce contexte, les effets liés au statut cognate s'expliquent de la manière suivante : lorsqu'un locuteur bilingue dénomme une image, les spécifications sémantiques activent simultanément deux représentations lexicales : la représentation lexicale cible de la langue sélectionnée pour la production et celle de l'autre langue non sélectionnée (Costa, Miozzo & Caramazza, 1999 ; Green, 1998) ; ces deux représentations activent ensuite simultanément

leurs phonèmes constitutifs. En conséquence, les phonèmes partagés reçoivent un supplément d'activation ; d'où des effets de facilitation. Par opposition, avec une image non-cognate, comme il n'y a pas de recouvrement phonologique, chaque phonème n'est activé que par une seule source d'activation. Ainsi, dans l'exemple cité plus haut, pour un locuteur bilingue espagnol-catalan dénommant l'image (cognate) d'un chat en espagnol, les représentations lexicales *gato* (espagnol) et *gat* (catalan) sont simultanément actives. Les phonèmes cibles /g/, /a/ et /t/ reçoivent donc de l'activation de *gato* et de *gat*. Dans le cas de l'image (non-cognate) d'un chien, comme la représentation lexicale *perro* (espagnol) et celle *gos* (catalan) ne partagent aucun phonème, les phonèmes cibles /p/, /e/, /r/ et /o/ ne sont activés que par l'item *PERRO*. La phonologie des images cognates est donc activée plus rapidement que celle des images non-cognates. Une transmission en cascade de l'activation rend compte des effets de facilitation liés au statut cognate des images sur les vitesses de dénomination chez des adultes bilingues (Kroll *et al.*, 1999 ; Costa *et al.*, 2000 ; Christofels *et al.*, 2008 ; Hoshino & Kroll, 2008). Ces résultats suggèrent que le même répertoire segmental est utilisé dans la production de l'une et l'autre langue (voir Roelofs, 2003 pour d'autres données en faveur de cette hypothèse).

IV - CONCEPTIONS INTERACTIVES DE LA PRODUCTION : DE QUELQUES ARGUMENTS

Comme déjà expliqué, une conception interactive de la production verbale est une conception en cascade qui inclut des rétroactions d'un niveau de traitement N +1 à un niveau N (Figure 1). L'interactivité peut être générale, c'est-à-dire que tous les niveaux de traitement sont reliés entre eux par des connexions bidirectionnelles ou, au contraire, elle peut être restreinte à certains niveaux de traitement. Le travail de simulation de Rapp et Goldrick (2000), déjà mentionné, a montré que seule une conception en cascade avec interactivité limitée entre les niveaux lexical et phonologique rendait correctement compte de certains types d'erreurs observées chez des patients. Il existe des données empiriques qui favorisent une telle conception comme nous allons l'exposer. Parmi les modèles de production qui comportent des liens bidirectionnels, il y a ceux de Dell et ses collègues (Dell, 1986, 1988, 1990 ; Dell & O'Seaghdha, 1991, 1992 ; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997) ou encore Cutting et Ferreira (1999).

Les erreurs mixtes

Les erreurs mixtes correspondent au fait que des erreurs sémantiques présentent une ressemblance phonologique avec leurs cibles comme par exemple la production par erreur de « cheval » pour « chien ». Ce phénomène a été avancé comme argument à l'encontre d'une conception discrète-sérielle, et donc, comme favorisant une conception interactive, comme celle de Dell et collaborateurs. Dans une conception interactive, les erreurs mixtes s'expliquent par les rétroactions (feedbacks) qui existent entre représentations lexicales (L) et phonologiques. Ainsi, l'activation du concept lexical CHIEN va se propager aux concepts CHAT, CHEVAL, RENARD, COCHON du fait de connexions qui existent à ce niveau. Les concepts activés transmettent alors de l'activation aux représentations lexicales correspondantes « chat », « cheval », « renard ». Enfin, ces unités lexicales activent leurs représentations phonologiques (les phonèmes individuels associés). Un item lexical comme « chat » ou « cheval » a une probabilité plus forte d'être sélectionné par erreur qu'un item comme « renard » ou « cochon » (toujours pour le concept cible CHIEN) du fait qu'il reçoit de l'activation à la fois du niveau conceptuel et, par rétroaction, du niveau phonologique : les phonèmes de l'item cible CHIEN renvoient de l'activation à tous les items lexicaux qui contiennent ces phonèmes. L'item lexical « renard », lui, ne reçoit de l'activation qu'en provenance du niveau conceptuel.

La conception discrète-sérielle, telle qu'implémentée par Roelofs (1992, 1997), semble exclure l'occurrence d'erreurs mixtes. En effet, selon cette conception, seul le lemma sélectionné est encodé phonologiquement, les erreurs mixtes ne peuvent donc se produire. Toutefois, les tenants d'une conception discrète-sérielle ont avancé que ces erreurs pouvaient s'expliquer de la façon suivante : dans des circonstances exceptionnelles, deux lemmas peuvent être sélectionnés et être encodés phonologiquement. Les cas de sélection double correspondraient donc à des occurrences rares de transmission en cascade (Levelt *et al.*, 1999). Aussi Levelt *et al.* font-ils appel à une instance de contrôle du langage (un éditeur post-lexical) pour expliquer ces erreurs. Les erreurs qui sont similaires sémantiquement et phonologiquement (les erreurs mixtes) auraient une probabilité moins élevée d'être détectées par l'éditeur, d'où leur plus grande incidence par rapport à d'autres types d'erreurs.

Les biais lexicaux

Dans une situation de communication spontanée, la production d'erreurs phonologiques, par exemple la production de l'expression « fesse molle »

au lieu de « messe folle » (tout le monde reconnaît dans l'exemple le clin d'œil à Rabelais) par des adultes est relativement rare (moins de 4 pour 10,000 ; Bonin, 2007). Ces erreurs aboutissent plus fréquemment (avec une probabilité supérieure au hasard) à la production d'un mot que d'un non-mot (par ex. Dell & Reich, 1981 ; Hartsuiker, Anton-Mendez, Roelstraete, & Costa, 2006 ; Nootboom, 2005a). Cette observation a été dénommée : « biais lexical ». Elle a été relevée dans des corpus d'erreurs et a aussi été confirmée par des études de laboratoire conduites dans différentes langues (par ex. en anglais : Baars, Motley, & MacKay, 1975 ; Dell, 1986, 1990 ; Motley, 1980 ; Hartsuiker, Corley, & Martensen, 2005 ; Humphreys, 2002 ; en espagnol : Hartsuiker *et al.*, 2006 ; Costa, Roelstraete, & Hartsuiker, 2006 ; en allemand : Nootboom, 2005b ; Nootboom & Quené, 2008). Les biais lexicaux ont été interprétés comme étant en faveur de l'existence de liens rétroactifs dans les modèles de production. Nous allons expliquer pourquoi plus loin. Il faut toutefois considérer avant les quelques études ayant remis en question ce phénomène.

Une remise en question de ce type d'erreurs a été argumentée sur la base de l'analyse d'erreurs phonologiques en anglais (Garrett, 1975) et en espagnol (Del Viso, Igoa, & Garcia-Albea, 1991). En effet, ces chercheurs ont rapporté des taux d'erreurs phonologiques se soldant par des biais lexicaux dont la probabilité d'occurrence ne dépassait pas celle du hasard. La probabilité qu'une erreur phonologique donne lieu par hasard à un mot réel avait été calculée sur la base de substitutions aléatoires d'un phonème à différentes positions à l'intérieur des mots. Toutefois, pour Dell et Reich (1981) et Hartsuiker *et al.* (2006), une telle estimation du hasard pose problème. Pour ces chercheurs, elle ne tiendrait pas suffisamment compte du voisinage phonologique des mots, c'est-à-dire le nombre de mots de même longueur qui ne diffèrent que par un phonème à une certaine position (par ex. Luce & Pisoni, 1998). Par exemple, la substitution aléatoire du premier phonème du mot « boule » (c'est-à-dire /b/) a une probabilité plus élevée de donner lieu à un mot réel (par ex. foule, poule, moule, roule, saoule) que la substitution aléatoire du troisième phonème (c'est-à-dire /l/ ; par ex. bouche, bouffe) car le voisinage du mot « boule » est plus important sur le premier phonème que sur le troisième. De plus, la substitution aléatoire d'un phonème du mot « boule » a une probabilité plus forte de conduire à un mot réel qu'une substitution du même type dans le mot « berge », car le premier possède davantage de voisins phonologiques que le second. Ainsi, en tenant compte dans leurs estimations du nombre de voisins phonologiques des mots et de leur distribution, Dell et Reich (1981) et Hartsuiker *et al.* (2006) ont observé un biais lexical dans les erreurs

phonologiques en anglais et en espagnol contrairement aux affirmations de Garrett (1975) et Del Viso *et al.* (1991).

Le biais lexical a été principalement mis en évidence dans des expériences menées en laboratoire. Plus précisément, un paradigme d'induction d'erreurs a été utilisé : le paradigme dit SLIP (Baars & Motley, 1974). Dans cette technique, les participants lisent d'abord silencieusement une série de paires de mots (ou de non-mots) – les paires « amorces » – qui sont présentées rapidement sur un écran d'ordinateur. Ces paires amorces sont ensuite suivies d'une paire de mots cibles que le participant doit à lire à voix haute à partir d'un signal. Pour illustrer, considérons les paires amorces suivantes : « fille-banc », « folle-bulle », « fort-bille ». Ces paires doivent rendre la succession des phonèmes /f/ et /b/ saillante. Si la paire de mots cibles est « boule-fiche », les amorces incitent à la production de l'erreur « foule biche ». Dans cet exemple, l'inversion des phonèmes initiaux se traduit donc par la production erronée d'une paire de mots réels. Au contraire, la paire de mots cibles « balle-film » précédée de la même série d'amorces doit conduire à la production de l'erreur « falle bilm ». Ici, l'inversion des phonèmes initiaux donne donc lieu à la production erronée d'une paire de non-mots. Grâce à ce paradigme, Baars *et al.* (1975) ont mis en évidence un biais lexical, de sorte que les erreurs phonologiques consistaient plus fréquemment en des mots réels qu'en des non-mots. L'interprétation de ce phénomène joue un rôle majeur dans l'opposition entre les tenants d'une conception discrète-sérielle (par ex. Levelt, 1989 ; Levelt *et al.*, 1999 ; Roelofs, 2004) et ceux d'une conception interactive (par ex. Stemberger, 1985 ; Dell, 1986) de la production verbale orale de mots.

Le biais lexical a été interprété en faveur de l'existence de liens rétroactifs (par ex. Dell, 1986 ; Rapp & Goldrick, 2000). Les erreurs formelles dans une conception interactive se produisent en raison du fait qu'un item lexical incorrect est sélectionné (par ex., « char » au lieu de « chat ») et ensuite correctement phonologiquement encodé ou bien parce que les phonèmes individuels ont été mal sélectionnés alors que l'item lexical correct a été choisi. Ainsi, un phonème anormalement activé, peut renforcer l'activation d'autres représentations lexicales, par le jeu de liens rétroactifs. Or, les non-mots n'étant pas représentés dans le lexique mental, ils ne peuvent bénéficier d'une telle activation additionnelle. Ainsi, un phonème anormalement activé risque-t-il donc davantage d'engendrer un mot réel qu'un non-mot.

Cependant, ce même phénomène est compatible avec une architecture discrète-sérielle qui inclut un mécanisme de contrôle du langage (« monitoring » en anglais : Baars *et al.*, 1975 ; Humphreys, 2002 ; Motley, 1980 ; Levelt, 1989 ; Levelt *et al.*, 1999 ; Nootboom, 2005b ; Nootboom

& Quené, 2008). Selon Levelt (1983, 1989), il existe deux voies de contrôle du langage – une interne et une externe – qui opèrent via le système de compréhension du langage. La voie externe nécessite d'entendre son propre discours tandis que celle interne, qui est supposée être à la source des biais lexicaux, met en jeu l'évaluation du discours planifié en envoyant en retour un mot phonologique généré au système de compréhension. Comme l'exprime Roelofs (2004), il y a donc une rétroaction de l'activation des sons aux items lexicaux dans la conception de Levelt *et al.* (1999) sauf que le *feedback* engage le système de compréhension du discours et non le système de production du langage.

Selon Levelt (1989 ; Levelt *et al.*, 1999), le mécanisme de contrôle évalue les erreurs potentielles en s'appuyant sur un critère de lexicalité (c'est-à-dire « les segments prêts à être articulés forment-ils un mot réel ? »). Le cas échéant, il détecte l'erreur et la corrige en réinitialisant le processus de production. Les erreurs correspondant à des non-mots (par ex. « balle-film » ==> « falle bilm ») seraient plus facilement détectées et rejetées par ce mécanisme que celles correspondant à des mots réels (par ex. « boule-fiche » -> « foule biche »). Les biais lexicaux peuvent donc s'interpréter sans faire appel à des liens rétroactifs entre nœuds lexicaux et phonèmes individuels.

Plusieurs résultats semblent en faveur d'une interprétation des biais lexicaux en terme de contrôle du langage, et donc, ils ne peuvent a priori servir d'arguments à une conception interactive. Avec la technique d'induction d'erreurs SLIP, Baars *et al.* (1975) ont fait varier le statut lexical des paires d'amorces et des cibles en anglais. Dans une expérience, les paires amorces et cibles correspondaient à des mots. Dans une autre, la liste était composée d'amorces et de cibles non-mots. Les échanges de phonèmes initiaux entre les non-mots des paires cibles conduisait soit à un mot réel, soit à un non-mot. Par exemple, les paires de non-mots « bire-fape » et « bane-foume » amorcent la production des mots « bulle fort » au lieu de celle cible « fulle-bort » (ou bien les non-mots « bime fure » à la place de la cible « fime-bure »). Enfin, dans cette même expérience, une autre liste contenait des paires-amorces et des paires-cibles qui étaient, tantôt des non-mots, tantôt des mots réels (c'est-à-dire contexte de listes mixtes). Les résultats de Baars *et al.* (1975) ont montré que dans un contexte de mots, comme dans un contexte mixte, les erreurs commises par les participants étaient plus fréquemment des mots que des non-mots : il y avait donc un biais lexical. En revanche, lorsque les paires amorces et les paires cibles correspondaient exclusivement à des non-mots, les erreurs donnant lieu à des mots réels n'étaient pas plus nombreuses que celles donnant lieu à des non-mots. Ce patron de résultats a été pris comme argument en faveur de l'existence d'un mécanisme de contrôle qui serait

sensible au contexte de production, contexte qui serait à l'origine du biais lexical (Levelt, 1989 ; Levelt *et al.*, 1999 ; Nootboom, 2005 ; Roelofs, 2004). En effet, dans un contexte où il n'y a que des mots, le mécanisme de contrôle s'appuierait sur un critère lexical (c'est-à-dire « est-ce un mot ? ») afin de détecter les erreurs. Selon Baars *et al.* (1975), le critère de lexicalité serait maintenu dans un contexte mixte. Cependant, dans un contexte constitué uniquement de non-mots, le critère de lexicalité serait abandonné, d'où la part égale d'erreurs lexicales et non-lexicales détectées (voir aussi Nootboom, 2005b). La sensibilité à la lexicalité dans les erreurs de production pose un sérieux problème à ceux qui l'ont interprété comme argument en faveur de l'existence de liens rétroactifs dans la production verbale. En effet, selon cette dernière interprétation, quelque soit le contexte de production les erreurs correspondant à des mots devraient être plus fréquentes que celles correspondant à des non-mots car le mécanisme interne de *feedback* se déroule de façon automatique et irrépessible.

Un autre résultat qui favorise une interprétation non-interactive du biais lexical a été obtenu à l'aide du paradigme SLIP. Nootboom (2005b) a montré que les biais lexicaux étaient fonction de la distance phonétique entre erreur et cible. La distance phonétique renvoie au fait que certains phonèmes sont plus proches phonétiquement que d'autres (par ex. le phonème /l/ est plus proche de /r/ que de /b/). Ainsi, les paires cibles « roue-biche » et « route-bol » ont des phonèmes initiaux, /r/ et /b/, phonétiquement éloignés alors que pour les paires « rond-lampe » et « ring-luge », /r/ et /l/ ont des phonèmes phonétiquement proches. Nootboom (2005b) a fait varier la distance phonétique entre les phonèmes des paires de mots cibles et ceux des erreurs induites. Ses résultats attestent d'un biais lexical, de sorte que les erreurs produites sont plus fréquemment des mots que des non-mots. Toutefois, le biais lexical est plus faible lorsque les erreurs produites sont phonétiquement proches des cibles que lorsqu'elles sont phonétiquement distantes. Pour Nootboom (2005b), cette interaction entre le biais lexical et la distance phonétique reflète la difficulté du mécanisme de contrôle à détecter les erreurs correspondant à non-mots trop proches phonétiquement des paires cibles. Le fait que la distance phonétique affecte la probabilité de détection des non-mots erronés l'amène à considérer, comme Levelt (1989 ; Levelt *et al.* 1999), que le contrôle de la parole interne fait appel au même mécanisme que celui qui est utilisé dans la reconnaissance des mots parlés (sur cet aspect voir aussi Özdemir, Roelofs & Levelt, 2007). Les résultats de Nootboom (2005b) sont donc, là encore, problématiques pour une interprétation du biais lexical favorisant une architecture interactive qui ne comporte pas de mécanisme de contrôle du langage.

Cependant, le problème des biais lexicaux n'est pas entièrement réglé. Costa *et al.* (2006) ont montré que, non seulement un biais lexical survient chez des bilingues qui produisent dans leur langue seconde, mais également, que le biais lexical n'est pas restreint à la langue de production. Des bilingues espagnol-catalan lisaient silencieusement une série de paire de mots-amorces en espagnol, puis à voix haute une paire de mots cibles espagnols. Un échange accidentel des phonèmes initiaux au sein des paires cibles résultait soit en une paire de mots catalans dans une condition (par ex. « *tip-nas* » > « *nip tas* »), soit en une paire de non-mots dans l'autre (par ex. « *til-naf* » > « *nil taf* »). Leurs résultats ont attesté que les erreurs se soldaient par une production accrue de mots catalans en comparaison à celle de non-mots (un biais lexical est observé), alors même que la tâche nécessite une production en langue espagnole ! Pour Costa *et al.* (2006), ces résultats favorisent une conception interactive dans laquelle le mécanisme de feedback serait à l'origine du biais lexical. L'existence d'un biais lexical interlangue est difficilement réconciliable avec une conception discrète-sérielle comportant un mécanisme de contrôle. Les résultats de Costa *et al.* (2006) s'expliquent aisément dans une conception interactive de la production bilingue dans laquelle existent des rétroactions entre phonèmes individuels et représentations lexicales.

Des données non réconciliables sur le biais lexical ?

Si l'étude princeps de Baars *et al.* (1975) est en faveur d'une interprétation du biais lexical en terme de mécanisme de contrôle, et donc est compatible avec une architecture sérielle et discrète (Levelt, 1989), d'autres données récentes remettent en question cette interprétation comme celles de Hartsuiker *et al.* (2005a). Selon Hartsuiker *et al.* (2005a), l'hypothèse d'un mécanisme de contrôle « adaptatif » à l'origine du biais lexical prédit un biais lexical *inversé* dans un contexte de non-mot et sa disparition dans un contexte mixte. *A contrario*, une hypothèse en terme de *feedback* du biais lexical prédit un biais lexical quelque soit le contexte. Toujours à l'aide du paradigme SLIP, Hartsuiker *et al.* (2005a) ont observé que les biais lexicaux étaient modulés par le contexte de production. Comme Baars *et al.* (1975), ils ont observé un biais lexical dans un contexte mixte mais pas dans un contexte de non-mots. Toutefois, Hartsuiker *et al.* (2005b ; Expérience 2) ont observé que le nombre d'erreurs donnant lieu à des non-mots ne différait pas significativement entre contextes mixte et de non-mots. Cette observation suggère que, dans un contexte mixte, les erreurs aboutissant à des non-mots n'ont pas été supprimées par le mécanisme de contrôle comme le suggéraient Baars *et al.* (1975). Par ailleurs, une absence de biais

lexical était observée en condition de non-mots. L'étude de Hartsuiker *et al.* (2005a) suggère donc qu'une interprétation en terme de contrôle ne peut pas rendre compte de la présence d'un biais lexical dans un contexte mixte, et qu'une interprétation en terme de *feedback* ne parvient pas, non plus, à expliquer l'absence de biais lexical en contexte de non-mots. Autrement dit, aucune des deux interprétations ne parvient à rendre compte de manière satisfaisante de l'ensemble des résultats.

Si le phénomène du biais lexical a pu être interprété en faveur de deux conceptions théoriques opposées, comme l'ont souligné Hartsuiker *et al.* (2005a ; voir également Nootboom, 2005b ; Nootboom & Quené, 2008), les interprétations de ce phénomène ne sont pas mutuellement exclusives. Dans un contexte mixte, le mécanisme de contrôle ne pouvant filtrer les erreurs sur la base de leur statut lexical, serait inopérant ; et donc le *feedback* entre représentations lexicales et les phonèmes individuels serait responsable du biais lexical. Dans un contexte de non-mots, d'une part le *feedback* favoriserait la production de mots réels au détriment des non-mots ; d'autre part, le mécanisme de contrôle emploierait un critère *anti*-lexical, rejetant les mots comme étant des erreurs au profit des non-mots (voir Nootboom, 2005b). L'absence de biais lexical dans cette condition résulterait donc de la combinaison de ces deux effets antagonistes.

Comme nous venons de l'exposer, de nombreuses données empiriques qui mettent en évidence l'existence d'un biais lexical ont contribué à alimenter le débat entre partisans d'une conception discrète-sérielle, et ceux d'une conception interactive, en l'attribuant pour les premiers à un mécanisme de contrôle de la parole interne et à mécanisme de *feedback* pour les seconds. Aucune des deux interprétations n'est à même d'expliquer l'ensemble des résultats obtenus. Seule une architecture interactive disposant d'un mécanisme de contrôle rend compte de l'intégralité des données disponibles sur le biais lexical.

Données sur les images cognates

Comme nous l'avons présenté, les effets de facilitation du statut cognate, obtenus en dénomination orale d'image chez des locuteurs bilingues, ont été interprétés comme favorisant une architecture en cascade (par ex. Costa *et al.*, 2000 ; Costa, Santesteban, & Cano, 2005). Costa *et al.* (2006) ont aussi proposé que ces effets s'interprétaient en termes de connexions bidirectionnelles entre représentations lexicales et phonèmes individuels. Selon cette hypothèse, quand un locuteur bilingue dénomme une image cognate, la représentation lexicale non-cible c'est-à-dire celle qui

correspond à l'autre langue, renforce l'activation des phonèmes partagés avec la cible. Ces derniers renvoient de l'activation aux représentations lexicales avec lesquelles ils sont liés, renforçant par là l'activation de la représentation lexicale cible. Pour une image non-cognate, la représentation lexicale activée dans la langue non-cible ne peut activer les phonèmes de la cible. La représentation lexicale correspondant à une image à dénommer est d'autant plus rapidement activée qu'elle bénéficie de nombreuses rétroactions en provenance du niveau des phonèmes, comme c'est le cas avec des stimuli cognates. Une architecture interactive permet donc également d'expliquer les effets facilitateurs des images cognates sur les latences observées dans différentes expériences de dénomination en temps réel chez des bilingues (Costa *et al.*, 2000 ; Janssen, 1999 ; Kroll *et al.*, 1999 ; Hoshino & Kroll, 2008 ; Christofels *et al.*, 2007).

Les résultats de Gollan et Acesnas (2004) obtenus dans une tâche d'induction de MBL s'accordent aussi avec l'hypothèse de liens bidirectionnels en production verbale. Ces chercheurs ont observé que des bilingues étaient moins sujets à des états de MBL lorsqu'ils dénommaient des images cognates que des non-cognates. Cette observation suggère que l'activation phonologique de la représentation lexicale de la langue non sélectionnée favorise l'activation de la représentation lexicale cible de la langue de production sélectionnée. L'hypothèse d'un mécanisme de *feedback* permet de rendre compte de l'influence des unités phonologiques durant la sélection lexicale².

Une architecture interactive de la production verbale peut rendre compte aussi de résultats obtenus à partir d'images cognates chez des bilingues aphasiques. En effet, la production de patients bilingues aphasiques est influencée par le statut cognate des images. Roberts et Deslauriers (1999) ont présenté à des personnes aphasiques et à des adultes sains, les uns et les autres étant bilingues français-anglais, des images cognates et non-cognates. De façon peu surprenante, les patients commettaient davantage d'erreurs que les adultes sains. Toutefois, les images cognates facilitaient la dénomination en comparaison de celles non-cognates, et ce, pour les personnes aphasiques comme celles non-aphasiques. Cela se traduisait par un taux d'erreurs de dénomination plus faible pour les images cognates que pour celles non-cognates. Ces résultats suggèrent que la phonologie de la langue non sélectionnée est activée, laquelle, lorsqu'elle partage des phonèmes avec la cible, permet

²Toutefois, un mécanisme de cascade sans *feedback* peut aussi bien expliquer les effets du statut cognate des images sur le taux d'occurrence de MBL si l'on considère que le MBL survient à une étape post-lexicale lors d'une difficulté d'accès aux phonèmes (Burke, Mackay, Worthley, & Wade, 1991 ; Gollan & Acesnas, 2004).

d'accéder plus facilement à la représentation lexicale cible. Konhert (2004) a également rapporté le cas d'un patient aphasique bilingue espagnol-anglais, dont la dénomination était influencée par le statut cognate des images. Le patient DJ présentait de faibles performances en dénomination d'images en espagnol comme en anglais, et dénommait les images cognates plus précisément que celles non-cognates. Afin d'améliorer sa performance en dénomination, DJ avait reçu un entraînement étalé sur plusieurs semaines. Une vingtaine d'images (10 images cognates et 10 images non-cognates) lui étaient régulièrement présentées pour qu'il les dénomme en espagnol. À l'issue de cet entraînement, sa performance de dénomination s'était améliorée pour les deux types d'images. De plus, sa performance en dénomination en anglais à partir des mêmes images cognates avait aussi progressé. Ces observations suggèrent donc un transfert positif de l'entraînement d'une langue à l'autre. Pour Costa *et al.* (2005), un tel transfert n'intervient que lorsque l'activation de la phonologie peut influencer la sélection lexicale. Comme nous l'avons donc exposé, une conception interactive de l'accès lexical chez des bilingues, rend compte de nombreux résultats expérimentaux sur l'impact du statut cognate des images à dénommer, et ce, chez des adultes sains comme chez des adultes aphasiques.

Données sur l'influence du voisinage en dénomination

Le voisinage d'un mot a été défini de plusieurs manières. Ainsi le voisinage orthographique d'un mot correspond-il aux mots qui peuvent être obtenus par substitution d'une lettre (Coltheart, Davelaart, Jonassen & Besner, 1977). Le mot « maison » selon cette définition a pour voisins orthographiques : « raison » et « saison ». Le voisinage phonologique est l'analogue du voisinage orthographique appliqué aux phonèmes (voir aussi Luce & Pisoni, 1998). De nombreuses études se sont penchées sur l'impact du voisinage en perception visuelle ou auditive des mots. En ce qui concerne la perception visuelle ou auditive de mots, de très nombreux résultats montrent que le nombre de voisins phonologiques (densité) ralentit le traitement lexical (en perception auditive, par ex. Vitevitch & Luce, 1998), soit au contraire le facilite (en perception visuelle, par ex. Yates, 2005). D'autres aspects du voisinage phonologique peuvent avoir une influence dans des tâches de perception, comme sa distribution (Vitevitch, 2007). Il n'est pas dans notre intention de fournir une vue d'ensemble de ces travaux en perception des mots (voir Spinelli & Ferrand, 2005, pour une synthèse). Ce qui nous intéresse ici, c'est l'impact du voisinage phonologique en production verbale conceptuellement dirigée. Les travaux sur ce thème en

dénomination ne sont pas nombreux. Pourtant, ils sont critiques en regard de la question de l'interactivité en production verbale.

Vitevitch (2002) a réalisé une expérience dans laquelle des participants devaient dénommer des images. Les noms des images pouvaient avoir un voisinage phonologique dense ou au contraire un voisinage restreint. Les résultats ont montré que la vitesse de dénomination était plus rapide pour des images ayant un voisinage élevé comparativement à celles ayant un voisinage plus faible. Ce résultat est clairement en accord avec une conception interactive de la production. Selon cette dernière, la facilitation liée au voisinage tient à ce que les représentations phonologiques des cibles ayant un voisinage élevé reçoivent plus d'activation que celles des cibles ayant un voisinage moins élevé. Plus précisément, comme l'illustre la Figure 1, la forme lexicale d'un mot active les nœuds phonologiques correspondants. À leur tour, ces nœuds phonèmes vont renvoyer de l'activation au niveau des formes lexicales, créant par là, un cycle de renforcement (positif) de l'activation. Comme nous l'avons déjà dit, dans une telle architecture, les phonèmes des mots ayant un voisinage dense reçoivent donc plus d'activation de la part du niveau mot que les phonèmes des mots ayant un voisinage plus étroit (voir aussi Gaskell & Dumay, 2003 pour d'autres résultats en accord avec ceux de Vitevitch).

V - LA QUESTION DES TRANSFERTS D'ACTIVATION HORIZONTAUX

Afin de donner une vue d'ensemble sur la circulation de l'information en production verbale, nous évoquons brièvement la question des transferts d'activation horizontaux. On s'intéresse ici à la planification de la réponse verbale lorsque plusieurs unités linguistiques doivent être agencées linéairement, c'est-à-dire lorsque des syntagmes, des phrases etc., et non plus seulement un mot isolé, doivent être produits. Notre objectif est toutefois modeste. Il s'agit seulement ici de sensibiliser les lecteurs aux travaux qui traitent de cette question. Les études conduites dans cette perspective visent à déterminer le degré de planification des unités à produire dans une séquence linguistique (par ex. un syntagme « une grosse voiture ») avant son initialisation. Selon les niveaux de traitement concernés, la planification peut être plus ou moins large. Certaines erreurs de production à l'oral renseignent sur l'empan des unités planifiées comme les échanges de mots et de segments. Les échanges de segments (« *fesse molle* » au lieu de « *messe folle* ») portent sur des éléments qui

appartiennent à catégories grammaticales différentes mais les segments échangés relèvent du même syntagme, tandis que pour les échanges de mots (« la *boîte* est dans la *lettre* » au lieu de « la *lettre* est dans la *boîte* »), ils appartiennent à des syntagmes différents mais sont de même catégorie grammaticale (Garrett, 1975). Ces catégories d'erreurs ont été prises comme argument en faveur de l'hypothèse selon laquelle il existe deux niveaux de traitement dont l'empan varie. Le niveau syntaxique correspond à l'encodage grammatical et a un empan de planification plus élevé que le niveau phonologique lors duquel les phonèmes individuels sont récupérés et assemblés. Ces dernières années, des études ont été conduites sur ce problème à l'aide de paradigme temps réel. Nous présentons ci-après rapidement quelques-unes de ces études.

Dans une recherche déjà ancienne, Meyer (1996) a étudié la production verbale orale d'expressions linguistiques à partir de la présentation visuelle de couples d'images. Des participants devaient produire des expressions telles « l'arbre et la table » ou bien « l'arbre est à côté de la table ». Ils dénommaient en débutant par l'image située à leur gauche. Les paires d'images étaient accompagnées de distracteurs auditifs. Ainsi, pour un couple d'images, le distracteur était soit sémantiquement relié à la cible de gauche, soit sémantiquement relié à celle de droite, soit non-relié. Les distracteurs pouvaient aussi être reliés phonologiquement à la cible de gauche (par ex. « arche »), à la cible de droite (par ex. « tarte »), ou non-reliés (par ex. « pluie »). Ils étaient présentés à différents SOAs (0 et -150 ms dans la première série d'expériences et 0 ms, +150 ms et +300 ms dans la seconde). Un effet d'interférence sémantique était observé pour les SOAs de 0 ms et -150 ms et pour les deux noms, de sorte que les latences d'initialisation étaient plus longues lorsque les distracteurs étaient sémantiquement reliés aux premiers ou aux seconds noms des images que lorsqu'ils étaient non-reliés. Les latences d'initialisation articulatoires étaient plus courtes avec des distracteurs reliés phonologiquement aux premiers noms à produire qu'avec des distracteurs non-reliés. Pour le second nom, un effet d'inhibition (non significatif) était observé de façon récurrente dans les expériences. Donc, un effet de facilitation phonologique était observé seulement pour le premier nom à produire (par ex. « arbre » pour l'exemple choisi).

Pour Meyer (1996), l'effet d'interférence sémantique se produit au niveau des lemmas et la facilitation phonologique au niveau des lexèmes phonologiques (Schriefers *et al.*, 1990 ; Levelt *et al.*, 1999). Ainsi les résultats ont-ils été interprétés comme traduisant le fait que les lemmas des deux noms cibles sont sélectionnés avant l'initialisation de la parole alors que l'encodage phonologique est relativement séquentiel (même si les formes

phonologiques des deux noms sont activées comme le suggère la tendance à l'inhibition phonologique observée sur le second nom). Bonin, Fayol et Malardier (2000) se sont inspirés de l'étude de Meyer (1996) pour étudier l'encodage de la forme des mots en dénomination écrite de paires de mots avec le paradigme de l'interférence et y ont observé des résultats similaires.

Concernant spécifiquement l'étendue de la planification phonologique en production verbale orale, Costa et Caramazza (2002) ont réalisé une étude sur des syntagmes. Des participants produisaient à partir d'images d'objets en couleur des syntagmes de structure Déterminant + Adjectif + Nom (par ex. *the green plane*). Ces cibles imagées étaient accompagnées de mots distracteurs qui étaient ou non phonologiquement reliés au nom. Un effet de facilitation phonologique était observé sur les latences suggérant ainsi que l'intégralité du syntagme était planifiée avant l'initiation de la production (voir également Damian & Dumay, 2007). Dans une autre étude, Schnur, Costa et Caramazza (2006) ont étudié la planification d'expressions plus longues telles que *The orange girl walks* et ont montré que, comparativement à des distracteurs mots non-reliés, des distracteurs phonologiquement reliés au verbe (par ex. *walnut*) accélèrent l'initialisation de telles expressions.

Afin d'étudier le degré de planification « horizontale », il est possible de déterminer dans quelle mesure les caractéristiques associées à des unités N, N +1, N +2 etc. (par ex. fréquence lexicale, longueur, familiarité conceptuelle...) se reflètent sur les latences d'initialisation lorsqu'elles doivent être produites successivement. Ainsi dans l'expression « une chaise est à côté d'un parapluie », on peut examiner si et quelles caractéristiques associées à l'unité « parapluie » ont une influence sur les latences d'initialisation, en plus de celles associées à l'unité produite en premier « une chaise ». Dans une expérience, Damian et Dumay (2007) ont montré que lorsque des participants dénommaient des objets en couleur avec des syntagmes de structure Adjectif + Nom, les latences d'initialisation étaient plus rapides lorsque l'adjectif et le nom partageaient leur phonème initial (par ex. *red rug*) que lorsqu'ils ne le partageaient pas (par ex. *red goat*). Ces résultats confirment ce qui a été observé par Costa et Caramazza (2002) à l'aide du paradigme de l'interférence. Ainsi, les locuteurs planifient phonologiquement l'intégralité du syntagme avant son initialisation. Dans une étude conduite en français (Malardier, Bonin & Fayol, 2004), nous avons, comme Damian et Dumay (2007), fait produire à l'écrit à des adultes, des couples de mots à partir d'images en noir et blanc. Les labels des images étaient reliés soit sémantiquement, soit associativement, soit phonologiquement ou bien étaient non-reliés. Nous avons observé

un effet d'interférence sémantique, un effet de facilitation associatif et pas d'effet significatif lié à la relation phonologique. Toutefois, Damian et Stadthagen-Gonzalez (2009) ont observé en production verbale écrite de syntagmes, des effets de facilitation phonologiques. Ainsi, dans une expérience, des participants produisaient des expressions syntagmatiques de type Adjectif + Nom (par ex. *red goat*) à partir d'images d'objets dont les contours étaient en couleur. L'adjectif de couleur était ou non relié phonologiquement au nom de l'objet (par ex. *blue bed* versus *red goat*). Les latences d'initialisation étaient plus courtes lorsque le nom et l'adjectif débutaient par le même phonème que lorsque les phonèmes initiaux étaient différents.

Bonin *et al.* (2006) ont fait dénommer des couples d'images de façon « nue » comme « pomme-maison » ou dans des expressions telles qu'« une pomme et une maison ». Les participants devaient toujours débiter par l'image de gauche. Des analyses de régression multiple ont été conduites sur les latences. Les prédicteurs correspondaient à des caractéristiques associées aux images d'objet (complexité visuelle, accord nom-image. . .) ou à leurs labels (âge d'acquisition, fréquence lexicale). Ces caractéristiques étaient relatives, soit au premier membre du couple, soit au second. Il a été observé que les prédicteurs significatifs des latences d'initialisation étaient ceux qui indexaient à la fois les niveaux structuraux, sémantiques et lexicaux pour l'item en position de gauche (c'est-à-dire celui devant être produit en premier), tandis que pour celui en position de droite, les prédicteurs significatifs étaient ceux qui indexaient des niveaux prélexicaux. Ce patron de résultats suggère donc qu'avant l'initialisation de la production d'un couple d'items, la quasi-intégralité des traitements relatifs au premier item est réalisée, tandis que seuls des traitements prélexicaux sont effectués en ce qui concerne le second.

Cet examen rapide de quelques travaux sur les transferts d'activation horizontaux montre que ceux-ci ne s'effectuent pas de la même manière selon les niveaux de traitement considérés. La planification au niveau conceptuelle apparaît plus large que celle au niveau phonologique. De même, pour un même niveau de traitement, l'empan de planification varie selon différents paramètres parmi lesquels le degré d'intégration entre entités (Damian & Stadthagen-Gonzalez, 2009) et la syntaxe (Janssen, Alario, & Caramazza, 2008 ; Janssen & Caramazza, 1999). Concernant le degré d'intégration entre unités, lorsque des syntagmes doivent être produits (intégration élevée), des effets de facilitation phonologiques sont observés. Tel n'est pas le cas lorsqu'il s'agit de noms d'objets non-reliés qui sont produits soit de façon nue, soit avec connecteur (par ex. « et ») (Bonin *et al.*, 2006). Aussi la syntaxe propre à la langue exerce-t-elle une contrainte sur l'activation des représentations phonologiques. Dans une

étude, Janssen *et al.* (2008) présentaient à des adultes des objets en couleur, et ils devaient dénommer soit le nom de l'objet (par ex. « râteau ») soit sa couleur (par ex. « rouge »). La ressemblance phonologique entre le nom de l'objet et sa couleur était manipulée. L'expérience a été conduite chez des parleurs francophones et anglophones.

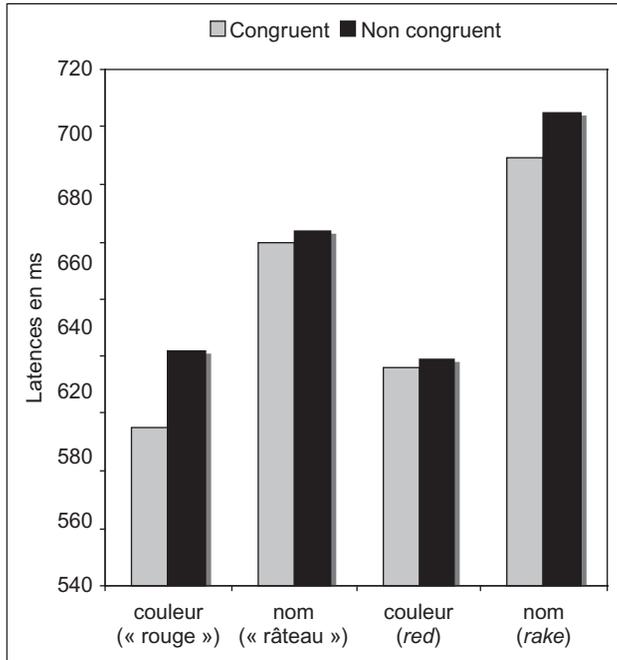


Figure 5. Latences en ms en fonction du type de tâche (dénomination du nom de l'image *versus* de sa couleur) et de la langue (français *versus* anglais).

Figure 5. Latencies in ms as a function of tasks (object naming *versus* color naming) and languages (French *versus* English).

Comme illustré par la Figure 5, pour les francophones, un effet de facilitation était observé lorsqu'ils dénommaient la couleur de l'objet (par ex. « rouge » pour l'image d'un râteau rouge) mais non lorsqu'il produisait son nom (« râteau »). Pour les anglophones, au contraire, un effet de facilitation phonologique était observé lorsqu'ils produisaient le nom de l'objet (par ex. *rake*) mais non lorsqu'ils énonçaient sa couleur (*red*). L'opposition des effets de facilitation phonologique s'explique par la position typique des adjectifs et des noms dans les deux langues :

« Nom + Adjectif » en français et « Adjectif + Nom » en anglais. Ainsi, en français, l'activation phonologique correspondant aux noms est initialement favorisée par rapport à celle des adjectifs et elle amorce la dénomination de l'adjectif lorsque les deux mots partagent le même phonème ; l'inverse est observé en anglais en raison du fait qu'il s'agit ici de la phonologie de l'adjectif qui est initialement favorisée.

Enfin pour terminer, des études suggèrent que le degré de planification horizontal n'est pas figé mais au contraire flexible. Ainsi, Schriefers et Teruel (1999) ont observé, à l'aide du paradigme de l'interférence, des effets de facilitation phonologique lors de la production de syntagmes nominaux en allemand (par ex. « table rouge ») avec des distracteurs qui étaient reliés phonologiquement à la syllabe initiale du premier mot alors qu'un effet plus faible de facilitation était observé avec des distracteurs reliés phonologiquement à la seconde syllabe du premier mot. Mais ces effets dépendaient toutefois des locuteurs considérés. Des analyses ont en effet permis de distinguer deux groupes de locuteurs. Pour le premier, les effets de facilitation phonologique étaient observés seulement avec des distracteurs reliés à la première syllabe du premier mot tandis que pour le second, les effets de facilitation étaient observés également avec des distracteurs reliés à la seconde syllabe du premier mot. Plus récemment, le caractère flexible de l'encodage grammatical a été mise en évidence par Wagner, Jescheniak et Schriefers (2010). Des participants devaient produire des phrases de différentes structures syntaxiques selon différentes conditions de charge cognitive. Quand ils décrivaient des arrangements correspondant à deux objets à partir de phrases simples telles que « la grenouille est à côté de la tasse », il a été observé que les deux noms étaient sémantiquement activés à un même degré avant l'initialisation de la réponse orale. Quand les adultes produisaient des phrases plus complexes comportant des adjectifs de couleur comme « la grenouille bleue est à côté de la tasse bleue », des effets d'interférence plus importants étaient observés sur le premier nom comparativement au second, suggérant par là que le second nom était sémantiquement activé avant l'initialisation de la réponse seulement dans un sous-ensemble des essais expérimentaux.

VI - CONCLUSIONS

De vifs débats en psycholinguistique cognitive de la production verbale ont porté sur la façon dont l'information se transmet d'un niveau de traitement à l'autre. Ces discussions ne sont pas aujourd'hui achevées.

Toutefois, comme nous espérons en avoir convaincu le lecteur, il est clair que les données disponibles sont en faveur au minimum d'une conception en cascade de la production verbale entre niveaux lexical holistique (L) et des phonèmes individuels. Il est intéressant à ce propos de mentionner, de façon anecdotique, que parmi les partisans les plus virulents d'une conception sérielle et discrète, certains admettent désormais clairement une transmission en cascade : c'est le cas d'Antje Meyer de l'université de Birmingham au Royaume-Uni qui avait cosigné en 1999, avec Levelt et Roelofs, l'article le plus abouti en faveur d'une théorie sérielle et discrète de la production verbale à l'oral.

Comme nous l'avons montré, la conception sérielle et discrète de la production verbale a été soutenue avec force et conviction dans de nombreuses publications. Comme déjà dit, selon nous les arguments en faveur de cette conception ont très souvent été « négatifs ». Ainsi comme nous l'avons expliqué, c'est par exemple l'absence d'effets significatifs qui a été prise comme argument en faveur de la non transmission de l'activation phonologique des voisins sémantiques coactivés à partir d'un concept cible (Levelt *et al.*, 1991). Face à des résultats en faveur d'une conception en cascade, les tenants de la conception discrète et sérielle ont soit proposé des interprétations ad hoc (par ex. Roelofs *et al.*, 1996), soit opposé des failles dans les démonstrations empiriques. Ainsi comme exposé, le fait que la fréquence lexicale interagisse avec la ressemblance structurale (Humphreys *et al.*, 1988) – résultat qui suggère très fortement une transmission en cascade de l'information en production – a été remis en cause en arguant d'une possible variable confondue de la fréquence lexicale avec la familiarité conceptuelle (Levelt *et al.*, 1999).

D'autres données doivent être recueillies pour renforcer une conception en cascade de la production verbale. En particulier, des études doivent être conduites sur ce thème en production verbale écrite. Il a souvent été admis que la production verbale écrite présentait une ressemblance fonctionnelle avec la production verbale à l'oral (Bonin & Fayol, 2000), cette première bénéficiant toutefois d'une relative autonomie (par ex. Bonin, Peereman, & Fayol, 2001 ; Rapp & Caramazza, 1997). Toutefois, on ne peut pas admettre a priori qu'un fonctionnement en cascade à l'oral signifie qu'un tel fonctionnement est valable à l'écrit. Des données récentes permettent d'argumenter que la transmission s'effectue aussi en cascade à l'écrit (Roux & Bonin, en préparation) du niveau lexical au niveau des graphèmes individuels.

Reçu le 11 décembre 2008.

Révision acceptée le 24 février 2010.

BIBLIOGRAPHIE

- Abdel Rahman, R., & Melinger, A. (2008). Enhanced phonological facilitation and traces of concurrent word form activation in speech production: An object naming study with multiple distractors. *Quarterly Journal of Experimental Psychology*, *61*, 1410-1440.
- Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382-391.
- Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access. *Journal of Memory and Language*, *48*, 468-488.
- Bonin, P. (2002). Les niveaux de traitement dans la production verbale orale et écrite de mots isolés à partir d'images. In M. Fayol (Ed.), *La production du langage*. Éditions Hermès.
- Bonin, P. (2007). *Psychologie du langage. Approche cognitive de la production verbale de mots*. Bruxelles : DeBoeck Université.
- Bonin, P., & Fayol, M. (2000). Writing words from pictures: What representations are activated and when? *Memory & Cognition*, *28*, 677-689.
- Bonin, P., Peereman, R., & Fayol, M. (2001). Do phonological codes constrain the selection of orthographic codes in written picture naming? *Journal of Memory and Language*, *45*, 688-720.
- Bonin, P., Roux, S., Barry, C., & Larell, L. (en préparation). Evidence for a limited-cascading account of written word naming.
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failure in young and older adults. *Journal of Memory and Language*, *30*, 542-579.
- Butterworth, B. L. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 108-135). Cambridge, London: MIT Press.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177-208.
- Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. *Cognition*, *64*, 309-343.
- Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192-208.
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance* (Vol. 6, pp. 535-555). New York: Academic Press.
- Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 1283-1296.
- Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, *41*, 365-397.
- Costa, A., Roelstraete, B., & Hartsuiker, R. (2006). The lexical bias effect in bilingual speech production: Evidence for feedback between lexical and sublexical levels across languages. *Psychonomic Bulletin & Review*, *13*, 612-617.
- Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, *94*, 94-103.

- Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 318-344.
- Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, 10, 111-117.
- Del Viso, S., Igoa, J. M., & García-Albea, J. E. (1991). Autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research*, 20, 161-185.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321.
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of prediction from a connectionist model. *Journal of Memory and Language*, 27, 124-142.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313-349.
- Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. *Psychological Review*, 4, 604-614.
- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287-314.
- Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20, 611-629.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and non aphasic speakers. *Psychological Review*, 104, 801-838.
- Ferrand, L. (1994). Accès au lexique et production de la parole : Un survol. *L'Année Psychologique*, 94, 295-312.
- Ferrand, L. (1997). La dénomination d'objets : Théories et données. *L'Année Psychologique*, 97, 113-146.
- Finkbeiner, M., & Caramazza, A. (sous presse). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex*.
- Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation*. Vol. 9 (pp. 133-175). New York: Academic Press.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production*. Vol. 1 (pp. 177-220). New York: Academic Press.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.
- Gaskell, M. G., & Dumay, N. (2003). Effects of vocabulary acquisition on lexical competition in speech perception and production. In M.J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th ICPHS Conference* (pp. 1485-1488). Adelaide, Australia: Causal Productions.
- Gollan, T. H., & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30, 246-269.
- Goldrick, M. (2006). Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes*, 21, 817-855.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language & Cognition*, 1, 67-81.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38, 313-338.

- Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related Reply to Baars, Motley, and MacKay (1975). *Journal of Memory and Language*, 52, 58-70.
- Hartsuiker, R. J., Antón-Méndez, I., Roelstraete, B., & Costa, A. (2006). Spoonish Spanerisms: A lexical bias effect in Spanish. *Journal of Experimental Psychology: Learning Memory and Cognition*, 32, 959-953.
- Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, 106, 501-511.
- Humphreys, G. W., Lamote, C., & Lloyd-Jones, T. J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory*, 3, 535-586.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, 5, 67-103.
- Humphreys, K. R. (2002). *Lexical bias in speech errors*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
- Janssen, N. (1999). *Bilingual word production: The time course of lexical activation in a mixed language context*. Unpublished Master's thesis, University of Nijmegen.
- Janssen, N., Schirm, W., Mahon, B., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34, 249-256.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological forms. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 824-843.
- Jescheniak, J. D., & Schriefers, H. (1997). Lexical access in speech production: Serial or cascaded processing? *Language and Cognitive Processes*, 12, 847-852.
- Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 1256-1274.
- Jescheniak, J. D., & Schriefers, H. (1999). Strictly discrete stages and contextual appropriateness. *Behavioral and Brain Sciences*, 22, 47-48.
- Jescheniak, J. D., Hahne, A., & Schriefers, H. (2003). Information flow in the mental lexicon during speech planning: Evidence from event-related brain potentials. *Cognitive Brain Research*, 15, 261-276.
- Jescheniak, J. D., Oppermann, F., Hantsch, A., Wagner, V., Mädebach, A., & Schriefers, H. (2009). Do perceived context pictures automatically activate their phonological code? *Experimental Psychology*, 56, 56-65.
- Knobel, M., Finkbeiner, M., & Caramazza, A. (2008). The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production. *Cognitive Neuropsychology*, 25, 256-286.
- Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language*, 91, 294-302.
- Kroll, J.F., Dijkstra, A.F.J., Janssen, N., & Schriefers, H.J. (1999). *Cross-language lexical activity during production: Evidence from cued picture naming*. In A. Vandierendonck, M. Brysbaert, & K. Van der Gooten (Eds.), 11th congress of the European Society for Cognitive Psychology. Proceedings (pp. 92). Gent: ESCOP / Academic Press.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122-142.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-38.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 1-36.
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, 33, 503-535.
- Malardier, N., Bonin, P., & Fayol, A. (2004). Dénomination par écrit de paires d'images : la question de la "dépendance". *L'Année psychologique*, 104, 407-432.
- Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory and Cognition*, 35, 494-503.
- Miozzo, M., & Caramazza, A. (1997). The retrieval of lexical-syntactic features in tip-of-the tongue states. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 1410-1423.
- Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General*, 132, 228-258.
- Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 555-63.
- Motley, M. T. (1980). Verification of "Freudian" slips and semantic pre-articulatory editing via laboratory-induced spoonerisms. In V.A Fromkin (Ed.), *Errors in Linguistic Performance. Slips of the Tongue (pp. 133-148), Ear, Pen, and Hand*. New York: Academic Press.
- Navarrete, N., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language*, 53, 359-377.
- Nickels, L. (2000). Spoken word production. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology* (pp. 291-320). Philadelphia, PA: Psychology Press.
- Nooteboom, S., (2005a). Listening to oneself: Monitoring speech production. In R. J. Hartsuiker, Y. Bastiaanse, A. Postma, & F.N.K. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 167-186). Hove, UK: Psychology Press.
- Nooteboom, S., (2005b). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication*, 47, 43-58.
- Nooteboom, S., & Quené, H. (2008). Self-monitoring and feedback: A new attempt to find the main cause of lexical bias in phonological speech errors. *Journal of Memory and Language*, 58, 837-861.
- Oppermann, F., Jescheniak, J. D., & Schriefers, H. (2008). Conceptual coherence affects phonological activation of context objects during object naming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34, 587-601.
- Özdemir, R., Roelofs, A., & Levelt, W. J. M. (2007). Perceptual uniqueness point effects in monitoring internal speech. *Cognition*, 105, 457-465.
- Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 539-557.
- Rapp, B., & Caramazza, A. (1997). From graphemes to abstract letter shapes: Levels

- of representation in written spelling. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1130-1152.
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, 107, 460-499.
- Roberts, P.M., & Deslauriers, L. (1999). Picture naming of cognate and non-cognate nouns in bilingual aphasia. *Journal of Communication Disorders*, 32, 1-14.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249-284.
- Roelofs, A. (2000). WEAVER ++ and other computational models of lemma retrieval and word-form encoding. In L. Wheeldon (Ed.), *Aspects of speech production* (pp. 71-114). Sussex (UK): Psychology Press.
- Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes*, 18, 175-204.
- Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review*, 111, 561-572.
- Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 246-251.
- Roux, S., & Bonin, P. (soumis). Cascaded processing in written naming: Evidence from the picture-picture interference paradigm.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time-course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Soni, M., Lambon Ralph, M. A., Noonan, K., Ehsan, S., Hodgson, C., & Woollams, A. M., (2009). "L" is for tiger: Effects of phonological (mis)cueing on picture naming in semantic aphasia. *Journal of Neurolinguistics*, 22, 538-547.
- Spinelli, E., & Ferrand, L. (2005). *Psychologie du langage : l'écrit et le parlé, du signal à la signification*. Paris : Armand Colin.
- Starreveld, P. A. (2000). On the interpretation of auditory context effects in word production. *Journal of Memory and Language*, 42, 497-525.
- Stemberger, J. P. (1985). An interaction activation model of language production. In A. Ellis. (Ed.), *Progress in the psychology of language* (pp. 143-183). Hillsdale (NJ) : Lawrence Erlbaum Associates.
- Sternberg, S. (1969). The discovery of processing stages: Extension of Donders' method. *Acta Psychologica*, 30, 276-315.
- Van Hell, J. G., & de Groot, A. M. B. (1998). Disentangling context availability and concreteness in lexical decision and word translation. *Quarterly Journal of Experimental Psychology*, 51A, 41-63.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science*, 9, 325-329.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 735-747.
- Vitevitch, M. S. (2007). The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition*, 35, 166-175.
- Vitkovitch, M., & Humphreys, G. W. (1991). Perseverative responding in speeded naming to pictures: Its in the links. *Journal of Experimental Psychology*:

Learning, Memory and Cognition, 17, 664-680.

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology:*

Learning, Memory and Cognition, 19, pp. 243-259.

Yates, M. (2005). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31, 1385-1397.