

# **FN5 : un nouveau modèle psycholinguistique informatique de la reconnaissance des mots dans la chaîne parlée du français**

*Nicolas Léwy et François Grosjean*

Laboratoire de traitement du langage et de la parole  
Université de Neuchâtel, Avenue du 1<sup>er</sup>-Mars 26, CH-2000 Neuchâtel, Suisse  
nicolas.lewy@unine.ch  
<http://www.unine.ch/ltlp/>

## **ABSTRACT**

We present a new computational psycholinguistic model (entitled FN5) of human spoken word recognition in French. The model contains a lexicon of 17,668 entries (in two versions, standard French and Swiss French), and it processes isolated words (determiners, prenominal adjectives, nouns) as well as sequences of two words (determiner + noun, prenominal adjective + noun). By means of a localist connectionist architecture with three levels (features, phonemes, and words), and a number of new characteristics (position processor, groups of connections, and isolation point), the model can simulate various phenomena (e.g. the recognition of words with a mute e or words linked with a liaison consonant). We end with a simulation example which illustrates the model's graphical user interface.

## **1. INTRODUCTION**

Cherchant à simuler sur ordinateur le traitement du langage et de la parole chez l'être humain, les modèles psycholinguistiques informatiques ont suscité beaucoup d'intérêt ces dernières années [1].

Jusqu'à présent, tous ceux proposés pour simuler la reconnaissance de mots parlés, comme TRACE [2] et SHORTLIST [3], se sont concentrés sur l'anglais. Le modèle que nous avons récemment élaboré, intitulé FN5, porte sur le français [4], et simule la reconnaissance de mots présentés en isolé (déterminant, adjectif antéposé, substantif) ou en suites de deux mots (déterminant et substantif, adjectif antéposé et substantif). Grâce à son interface utilisateur graphique, le modèle se prête bien à la démonstration.

## **2. LEXIQUE DU MODÈLE**

Le modèle contient un lexique de 16 971 substantifs, et 697 déterminants et adjectifs antéposés, pour un total de 17 668 mots, et ce dans deux versions, française et Suisse romande. (Dans la version Suisse romande, nous avons ajouté aux 35 phonèmes du français standard le /a/, le /œ/, et sept voyelles longues, /i: e: ε: ø: u: α:/, pour un total de 44 phonèmes en tout [5].)

Pour établir le lexique, nous avons extrait l'ensemble des déterminants, adjectifs et substantifs de BRULEX [6]

avant de les soumettre à une série de traitements à la fois informatiques et linguistiques. Nous avons éliminé les mots de certaines catégories (ex. substantifs au pluriel, adjectifs postposés uniquement, doublons, variantes orthographiques, mots de très basse fréquence, etc.), nous avons corrigé les indices de fréquence des mots à l'aide de LEXIQUE [7] (version 3) et les avons normalisés (les valeurs s'étendent désormais de 0 pour très rare à 1 pour très fréquent), et nous avons vérifié toute l'information linguistique présente dans le lexique (orthographe, transcription phonétique et genre). Nous avons également ajouté des informations sur les voyelles longues en finale de mot pour le français de Suisse romande (ex. *amie*, *vraie*, etc.) et sur les consonnes de liaison des adjectifs antéposés (ex. *grand* avec /t/, *heureux* avec /z/, etc.). Enfin, aux mots qui comportent un schwa (ex. *fenêtre*, *souvenir*, etc.), qu'il soit obligatoire dans la prononciation, facultatif ou interdit (2 190 mots en tout), nous avons associé des indices de préférence de la prononciation avec et sans effacement du schwa.

## **3. FONCTIONNEMENT DU MODÈLE**

Le modèle se sert du formalisme des réseaux connexionnistes localistes [8]. Dans la figure 1, nous présentons l'architecture générale du modèle qui s'inspire en partie des modèles antérieurs de la reconnaissance de mots parlés, notamment [2] et [3].

Le modèle consiste en un grand nombre d'unités de base qui travaillent toutes en parallèle. Ces unités sont organisées en trois niveaux linguistiques — traits distinctifs, phonèmes et mots —, et sont connectées par des liens d'activation (visualisés par les flèches pointues) et/ou d'inhibition (visualisés par les flèches rondes). Il existe des liens ascendants, des traits distinctifs vers les phonèmes ainsi que des phonèmes vers les mots, et des liens descendants, des mots vers les phonèmes (optionnels dans le modèle). De plus, il y a des liens latéraux, aux niveaux des phonèmes et des mots. L'entrée du modèle consiste en une suite de valeurs au niveau des traits distinctifs, représentant le mot (ou la suite de mots) à reconnaître. C'est en passant par les divers liens d'activation et d'inhibition, et au moyen d'un certain nombre de cycles de simulation, que les traits distinctifs répercutent leur état d'activation vers les phonèmes concernés, et que ces phonèmes interagissent à leur tour avec les mots qui les contiennent.

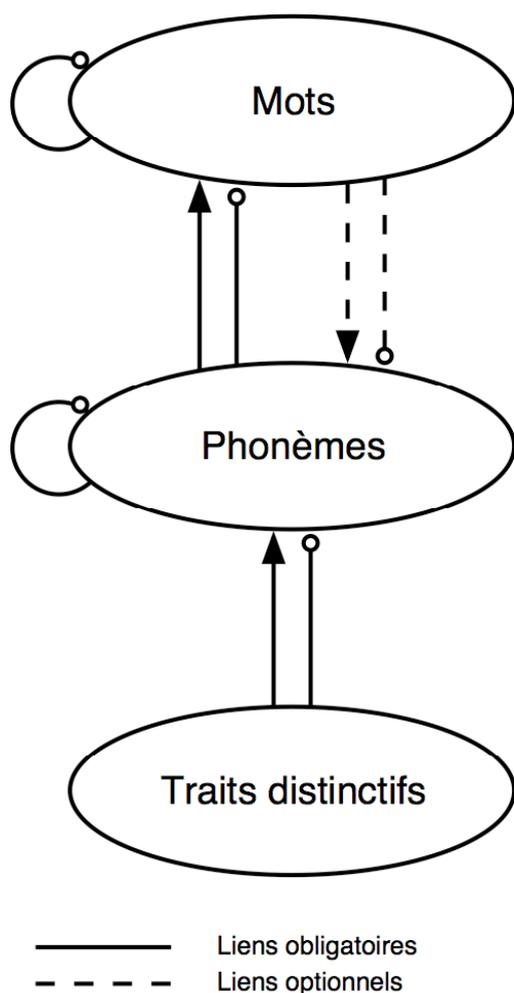


Figure 1 : Architecture générale du modèle.

Dans ce qui suit, nous présentons trois mécanismes développés spécifiquement pour le modèle.

### 3.1. Processeur de position

Afin de permettre la simulation de la reconnaissance d'une suite de deux mots (ex. *ta tasse* - /tatas/), nous avons prévu qu'à chaque cycle, un processeur de position aligne les mots du lexique avec chaque position possible de la chaîne des phonèmes en entrée. Dans le cas du mot *table*, par exemple, le modèle aligne le début du mot avec le début de la suite, à savoir, dans l'exemple /tatas/, avec le /t/ initial en position 1, mais également avec le /a/ en position 2, avec le deuxième /t/ en position 3, etc. Il établit pour chacune des positions la somme combinée des activations et inhibitions du mot, compare ces sommes, et choisit la position qui produit la somme la plus élevée pour ce mot. C'est cette position qui est alors retenue et stockée avec le mot pour la suite de la simulation. Ce positionnement optimal du mot est répété à chaque cycle d'un phonème (il y a 16 cycles par phonème) et cela pour chaque phonème de la chaîne phonétique. Certes, la position du mot peut changer d'un cycle à l'autre, mais nous imposons au système de n'en avoir qu'une seule, pour chaque mot, à chaque cycle. Si plusieurs positions donnent exactement le même résultat, le processeur de

position choisit la position la plus récente. Bien entendu, dans un traitement séquentiel où les phonèmes arrivent les uns après les autres (d'abord le premier phonème, ensuite le deuxième phonème, etc.), le processeur de position ne peut placer un mot que dans la position 1 lors de l'arrivée du premier phonème, et aucune comparaison n'est effectuée à ce moment-là. C'est à l'arrivée du deuxième phonème qu'il peut placer le mot en position 1 ou en position 2 en comparant les deux sommes, et c'est à l'arrivée du troisième phonème qu'il compare les sommes des positions 1, 2 et 3, etc. De la sorte, l'activation et la reconnaissance d'un mot peuvent avoir lieu à n'importe quel emplacement de la chaîne parlée.

Lorsque nous comparons cette approche avec celle d'autres modèles qui simulent la reconnaissance de deux ou plusieurs mots, comme [2] et [3], nous y voyons plusieurs avantages réunis en un seul modèle. Premièrement, l'approche avec processeur de position fonctionne sans qu'il faille dupliquer les mots à toutes les positions théoriquement possibles : les mots n'ont qu'une seule position qui leur est attribuée à chaque cycle et ils peuvent être déplacés vers une autre position si le processeur de position le décide. Deuxièmement, l'approche ne limite pas le nombre des mots-candidats qui sont pris en considération ; tous les mots du lexique peuvent être activés en parallèle, et ce avec un niveau d'activation approprié (forte lors d'une bonne correspondance, moindre lors d'une correspondance partielle, etc.). Troisièmement, l'approche ne nécessite pas que la chaîne en entrée soit segmentée préalablement ; elle s'appuie plutôt sur le fait que les frontières entre mots émergent souvent du processus normal de la reconnaissance des mots.

### 3.2. Groupements de connexions

Pour les mots qui contiennent des phonèmes qui peuvent être effacés (ex. un schwa) et/ou ajoutés (ex. une consonne de liaison), et plus généralement pour tous les mots avec plusieurs formes (comme les adjectifs), nous faisons appel aux groupements de connexions. Chacune des prononciations réalisables d'un mot (ex. avec ou sans schwa, avec ou sans consonne de liaison, masculin ou féminin, etc.) possède un groupement de connexions qui relie le mot avec les phonèmes en question. Par exemple, pour le substantif *pelouse*, il y a un groupement de connexions pour la forme avec effacement du schwa (/pluz/) et un autre pour la forme sans effacement (/pəluz/). De même, pour l'adjectif *grand*, il existe un groupement pour la forme féminine (/gʁɑ̃d/), un deuxième pour la forme masculine sans liaison (/gʁɑ̃/), et un troisième pour la forme masculine avec liaison (/gʁɑ̃t/).

Le processeur de position n'aligne donc pas seulement chaque mot avec chaque position possible dans la chaîne des phonèmes en entrée, mais aussi, dans le cas de plusieurs groupements de connexions, il établit les sommes des activations et inhibitions pour chaque groupement, à chaque position, afin de trouver le groupement et la position qui donnent la somme la plus

élevée (ex. le groupement qui correspond à la forme masculine avec consonne de liaison, et cela pour la position 2). A la fin de la comparaison, les informations concernant le groupement et la position sont stockées avec le mot. Le processus recommence à chaque nouveau cycle et peut donc produire un résultat différent, le processeur de position ne retenant qu'un seul groupement et qu'une seule position pour chaque mot. Notons que des informations numériques peuvent être attachées aux groupements de connexions afin de rendre compte de la préférence, ou de la fréquence, d'un groupement donné chez l'être humain (voir les mots avec schwa facultatif, par exemple, où il y a souvent une préférence pour l'une ou l'autre version).

### 3.3. Point d'isolation

Adaptée de la tâche psycholinguistique expérimentale du *gating* [9], nous utilisons une mesure de point d'isolation (PI) dans le modèle. Le PI est le moment, exprimé en nombre de cycles, où le mot à reconnaître dépasse le niveau d'activation de tous les autres candidats possibles ; de plus, ce premier rang doit être maintenu jusqu'à la fin de la simulation. Cette mesure est calculée pour chaque mot de la suite (dans le cas d'une suite de mots), et elle est proposée en nombre de cycles ainsi qu'en pourcentage de la longueur du mot (neutralisant ainsi la longueur). Le PI du modèle est une mesure fort utile dans la simulation de la reconnaissance de mots parlés.

## 4. EFFETS SIMULÉS

A l'aide de nombreux essais de simulation de la reconnaissance de mots, nous nous sommes assurés que le modèle peut rendre compte d'effets psycholinguistiques propres au mot en isolé (ex. les mots de haute fréquence prennent moins de temps à être reconnus que les mots de basse fréquence, les mots bisyllabiques prennent plus de temps que les mots monosyllabiques, etc.). En outre, le modèle simule la reconnaissance de suites de deux mots ; ceux-ci peuvent être séparés ou non par une pause, être liés par un enchaînement avec ou sans liaison (ex. *petit ami* et *chaque ours*), ou bien contenir un effacement du schwa (ex. *la s'maine*). Le modèle tient également compte de différentes prononciations régionales (ex. le fait que *ami* et *amie* soient homophones en français standard mais pas en français de Suisse romande).

## 5. EXEMPLE DE SIMULATION

Dans la figure 2, nous montrons la fenêtre principale de l'interface du modèle qui comporte deux parties. Celle du haut, avec divers boutons, zones de texte, etc., sert à régler les différentes composantes du modèle et à entrer les mots pour lesquels on veut simuler la reconnaissance ; celle du bas, libellée *Activation/Cycle*, sert à visualiser l'activation des mots.

A titre d'exemple, nous montrons la simulation de la reconnaissance de la suite *petit ami* (/ptitami/) prononcée avec un effacement du schwa dans *petit* et une liaison entre *petit* et *ami*. Pour obtenir le résultat présenté

dans la figure, nous avons d'abord chargé le lexique du modèle dans sa version française (voir le bouton *Lexique*, en haut à gauche). Ensuite, nous avons tapé les mots *petit ami* dans la zone de texte intitulée *Mot(s)*. La transcription phonétique /ptitami-/ est alors fournie par le logiciel dans la zone en haut, à droite du bouton *Lexique*, comme on le voit dans la figure. Notons que le /-/ marque la fin de la chaîne. Quant à la liaison, elle se manifeste visuellement de deux manières différentes : au niveau de la suite phonétique, un /t/ est inséré automatiquement au bon endroit, et dans la suite graphique (zone *Mot(s)*), un signe ~ apparaît entre *petit* et *ami* lorsque l'on fait un retour à la ligne après avoir tapé les mots. Finalement, nous avons cliqué sur le bouton *Entrer tout*, bouton qui sert à insérer les phonèmes dans le modèle, un par un, en commençant avec le premier phonème et en finissant par la marque de fin de suite. L'état d'activation des mots est alors apparu dans la zone *Activation/Cycle* de la figure, et a évolué, comme on le remarque en examinant la figure, au fur et à mesure de l'avancement dans la suite.

Certes, au tout début du processus (à gauche donc de la fenêtre), il y a de nombreux candidats possibles, d'où la masse de mots actifs. Mais on constate qu'au niveau du cycle 48 (à savoir à la fin des trois premiers phonèmes de la suite, /p/, /t/ et /i/), le mot *petit*, *-ite* est devenu le plus actif (voir la courbe blanche qui représente le niveau d'activation du mot entré). En fait, le point d'isolation (PI) de *petit*, *-ite* se trouve au cycle 36 déjà, durant le /i/ donc (voir le petit drapeau *PI* à gauche). Pendant l'entrée du /t/ de liaison (cycles 49–64), *petit*, *-ite* continue à monter tandis que des mots non compatibles (ex. les mots composés du type *petit-fils*) retombent. (Notons qu'en cliquant sur le bouton *6 1<sup>ers</sup>* (six premiers) ou *30 1<sup>ers</sup>* (trente premiers), il est possible d'obtenir une liste des candidats les plus actifs, et ce à tout moment de la simulation.) Quand le /a/ de *ami* arrive (cycles 65–80), quelques nouveaux candidats entrent en jeu (ex. *titane* et *italien*). Pendant l'entrée du /m/ (cycles 81–96), les mots en /am/ se renforcent (ex. *ami*, *amour*, *amiral*, etc. ; voir à droite le nuage de mots momentanément actifs). Enfin, c'est pendant la dernière voyelle et la marque de fin de suite (cycles 97–128) que *ami*, couplé à l'homophone *amie*, est isolé (PI au cycle 104). Le modèle vient donc de reconnaître la suite *petit ami*.

Notons que la chaîne en entrée /ptitami/ aurait aussi pu être reconnue par le modèle comme étant *petit tamis*, et donc sans liaison. En effet, *tamis* est assez actif à la fin de la simulation, mais comme ce mot est moins fréquent que *ami*, il se trouve à un niveau d'activation inférieur.

## REMERCIEMENTS

L'élaboration du modèle a été rendue possible grâce à deux subsides du Fonds national suisse de la recherche scientifique (FN 1214-058848.99 et 100012-103384). Lysiane Grosjean, Isabelle Racine et Carole Yersin ont participé à la préparation du lexique du modèle et à son évaluation ; qu'elles en soient remerciées. Nous tenons également à exprimer notre reconnaissance à Jeff Elman pour ses conseils très précieux tout au long du projet.

