

*The re-introduction of syllables into adult foreign language learning and pronunciation teaching : a creative challenge for learners and teachers based on typographical annotations, sound/ text synchronisation and greater self-awareness of quasi-permanent auditory perception problems called ‘mother tongue interference’*

**Conférence Internationale de CERCLES, Calabria, septembre 2016**

**Anthony Stenton, Lairdil, Université Toulouse 1**

# Introduction : the nature of Mother Tongue Interference (MTI)

Displaced lexical stress patterns can be a major source of linguistic confusion yet are rarely corrected by native speakers.

They can lead to ambiguity, communication problems, intolerance and **stigmatisation**.

Hypothesis : MTI is a linguistic variant of Cognitive Bias ( Kahnemann ) which requires specific pedagogical attention

# Intelligibility

Field (2005 : 417-418) : “Some EFL/ESL practitioners hold the view that perceptual errors are relatively trivial because listeners can compensate for them by drawing on information provided by context in the form of the listener’s understanding what has been said so far. This argument is circular, however, because context in this sense depends entirely on how much of the previous input the listener has been able to decode accurately.”

Field, J. (2005) “Intelligibility and the Listener: the Role of Lexical Stress”, TESOL Quarterly, 39(3), p. 399-423

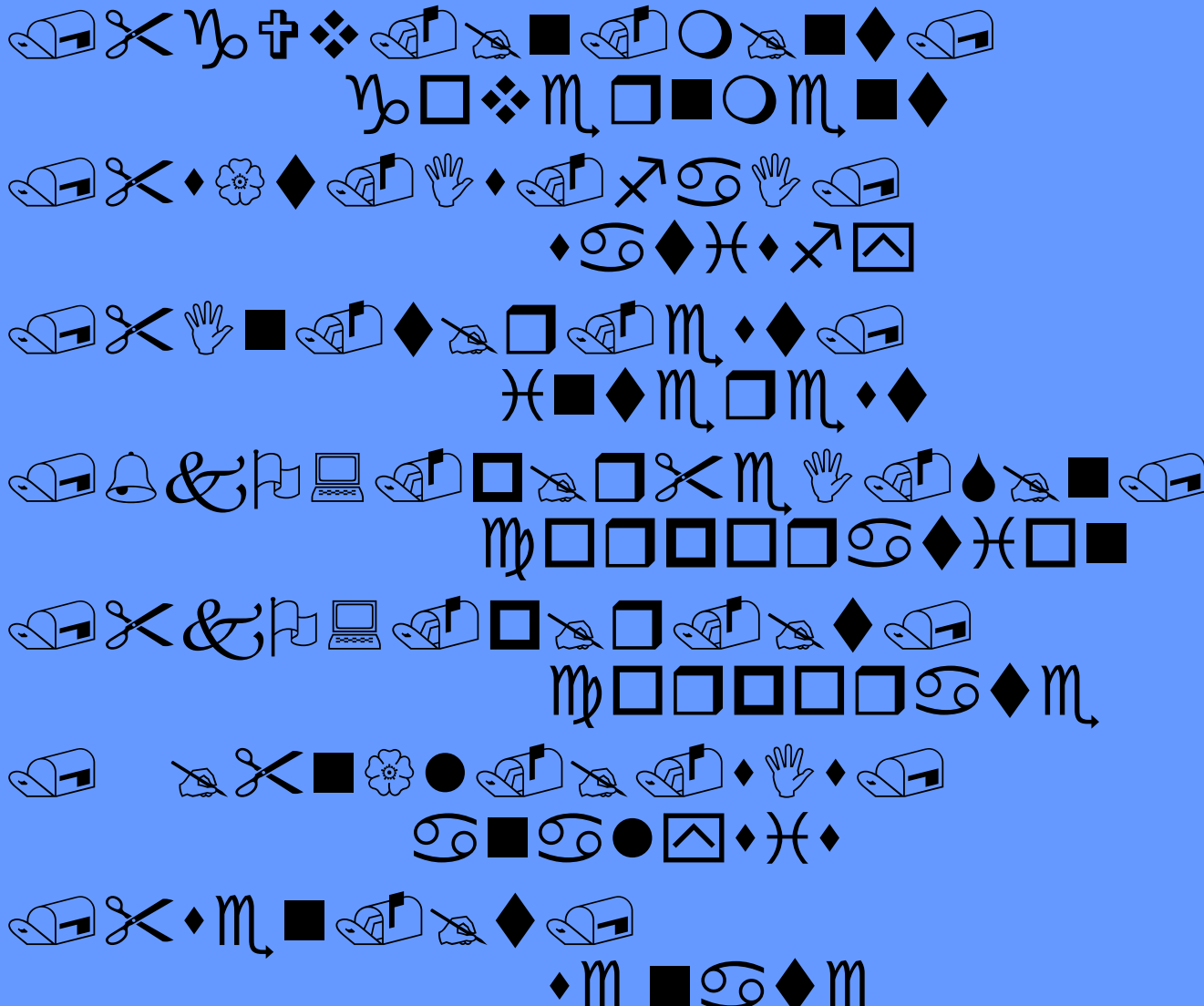
# Lexical stress and intelligibility

Lexical stress plays a central role in determining the profiles of words and phrases in theories of metrical phonology (Hogg & McCully, 1987). Psycholinguistic studies of slips of the ear (Bond, 1999) have shown that native English listeners place greater reliance on the stressed syllables of words than on the briefer unstressed syllables.

# L'interférence de la langue maternelle (1)

- L'interférence de la L1 inhibe la perception des schémas d'accentuation de la L2, ainsi que la perception des syllabes fortes et faibles.
- L'accentuation française est régulière et se fait sur la droite. On parle dans ce cas d'une langue « codachrone ».
- L'accentuation anglaise est irrégulière et se fait sur la gauche. On parle d'une langue « capochrone ».

Sur un total de plus de 250 étudiants de l'université de Toulouse, 70% à 90% se trompent de schémas d'accentuation:



Les problèmes d'interférence soulevés par les schémas d'accentuation lexicale anglais sont-ils les mêmes pour toutes les langues?

- Bien sûr que non
- En tchèque, letton, hongrois, allemand suisse, (dialecte Bernois), finlandais l'accentuation se situe toujours sur la première syllabe. (Le finlandais n'a pas de réduction de voyelle non accentuée.)
- En français, espagnol, portugais, turc et polonais, l'accentuation se fait rarement sur la première syllabe.

# L'interférence de la langue maternelle(2)

La qualité de l'accentuation varie selon les langues:

La hauteur (pitch)

La longueur

Le volume

Un phénomène complexe qui comporte plusieurs paramètres.



# Un phénomène universel

- L'interférence de la langue maternelle dans la perception et la production d'une L2 est un phénomène universel que l'on nomme «engagement neuronal », mais il est **amplifié** par l'alphabet LATIN, qui utilise les mêmes symboles pour décrire des systèmes de sons différents.

# Example of Italian speakers of German

Italian speakers tend to stress the second component of German morphologically complex words such as compounds and prefix verbs even if the first component is lexically stressed.

Some Italians who had lived 10 years in Germany never noticed this mistake and made it systematically. Clearly they did not perceive the L2 stress pattern.

Bissiri, Maria Paola and Pfitzinger, Hartmut R. (2009) *Italian speakers learn lexical stress of German morphologically complex words*. *Speech Communication*, Vol. 51 (10), p. 933-947. ISSN 0167-6393. Article.  
Italian

## Le problème de l'alphabet latin (3)

- La lecture et la sous-vocalisation (la lecture silencieuse) peuvent aggraver/renforcer les mauvaises habitudes. L'alphabet latin est un outil plurilingue développé et adapté en fonction de chaque langue par des mentalités **monolingues** et ethnocentriques.
- L'interférence est intimement liée à la proximité morphologique. Plus deux langues sont proches et plus l'interférence sera grande en raison du recours aux mêmes symboles latins pour décrire différents schémas d'accentuation lexicale.

# Le problème de l'alphabet latin

- Ne pas tenir compte des schémas d'accentuation lexicale revient à négliger/occulter les éléments acoustiques les plus essentiels de la chaîne de sons.
- L'accentuation s'accompagne d'une plus grande amplitude respiratoire.
- Hypothèse : Apprendre une L2 sans consciemment analyser le rôle des schémas d'accentuation lexicale pourrait tout simplement renforcer des problèmes d'interférence, ainsi «fossilisés » **à vie.**

# L'interférence de la langue maternelle

- Conclusion
- L'anglais et le français sont des langues morphologiquement amies, mais acoustiquement ennemies.
- 60% de la langue anglaise est composée de français ou de latin.
- Les schémas d'accentuation lexicale sont radicalement différents.

# Conséquences de l'interférence de la langue maternelle

- Rupture et échec de la communication
- Intolérance = les préjugés sont confortés et renforcés
- Stigmatisation
- 'Echec' comparatif (la France est 69ème sur 120 nations dans un classement mondial suite aux tests TOEFL de 2009)

# Conséquences de l'interférence de la langue maternelle (2)

- Les Anglais et les Français sont mutuellement démoralisés par leur incapacité respective à prononcer la L2 correctement. Ils sont victimes de l'alphabet latin, mais aussi d'une intolérance généralisée à l'égard de l'interférence de la langue maternelle.
- Au lieu de remettre en cause l'alphabet latin, ils se tiennent habituellement eux-mêmes pour responsables, invoquant des facteurs culturels (« perfide Albion ») ou blâmant un enseignement traditionnel jugé incompetent du fait de son parti pris littéraire.

# Conclusion

- La lecture d'une L2 AMPLIFIE l'interférence de la langue maternelle dans la perception et la production de la L2.



# Personalised learning?



cc Steve Wheeler, University of Plymouth, 2010

**anytime**

**personalised**

**anyplace**

Pour comprendre notre héritage culturel  
dans ce contexte il faut remonter aux  
sources

Peut-on imaginer que certaines activités  
liées à l'apprentissage d'une langue soient  
contre-indiquées, contre-productives?

Notre alphabet sert-il **nos** intérêts ?

Sert-il les intérêts des francophones ou les  
intérêts de l'homme ?

# L'alphabet phénicien prototype des alphabets modernes

L'alphabet phénicien est l'un des systèmes d'écriture les plus utilisés, transmis par les marchands phéniciens dans le monde méditerranéen où il a évolué et a été assimilé par de nombreuses cultures.

𐤀	𐤁	𐤂	𐤃	𐤄	𐤅	𐤆	𐤇	𐤈	𐤉	𐤊	𐤋	𐤌
aleph	beth	gimel	daleth	he	waw	zayin	heth	teth				
'	b	g	d	h	w	z	h	t				
𐤍	𐤎	𐤏	𐤐	𐤑	𐤒	𐤓	𐤔	𐤕	𐤖	𐤗	𐤘	𐤙
yod	kaph	lamed		mem	nun	samekh						
y	k	l		m	n	s						
𐤚	𐤛	𐤜	𐤝	𐤞	𐤟	𐤠	𐤡	𐤢	𐤣	𐤤	𐤥	𐤦
ayin	pe	sade	qoph	resh	shin							
'	p	s	q	r	sh/s							

# Aleph : symbole de tradition

*Aleph* est apparenté à l'alpha ( $A, \alpha$ ) de l'alphabet grec, au A ( $A, a$ ) de l'alphabet étrusque puis au A de l'alphabet latin et au А de l'alphabet cyrillique.



La tête du taureau avec ses 2 cornes est bien visible dans le caractère phénicien.

# Les enjeux

L'AL est un point d'entrée manifestement local et culturel pour l'apprentissage des langues européennes à la fois pour au niveau des L1 et des L2

Par conséquent, le cerveau européen est conditionné par ce système d'écriture (équipes de neuro linguistes : S. Dehaene et al.) et par ses **associations son/symbole**.

Par conséquent l'AL représente la plus grande source de confusion sonore et de confusion cérébrale sur la planète. Notre héritage culturel est bâclé. L'utilisation des mêmes symboles pour deux systèmes sonores différents est une façon peu élégante de se tirer une balle dans le pied.

# Enjeux

Il faut reconnaître le paradoxe : la lecture de la L2 est à la fois indispensable pour accélérer l'acquisition et nuisible au niveau de l'apprentissage de la perception auditive et de la production orale.

# Hypothèses

L'AL a été conçu par les monolingues pour les monolingues.

Personne n'a essayé à modifier les symboles de l'alphabet pour les besoins de précision multilingue avant le 19 siècle.

L'API (1888) et l'Espéranto (1887) sont des échecs au niveau de l'enseignement. Ils n'ont pas réussi à s'imposer pour l'apprentissage.

Il nous faut des symboles simples à mémoriser, faciles à apprendre et adaptés aux besoins multilingues du monde moderne.



# Hypothèses

- L'apprentissage d'une L2 est rendu difficile en grande partie par l'interférence à plus ou moins long terme de la L1 de chaque apprenant
- L'apprentissage de la lecture en L2 amplifie cette interférence. Notre perception auditive et production orale d'une L2 sont souvent déviantes
- L'interférence est universelle et durable et le résultat est une vaste ensemble de problèmes de communication et de **stigmatisation**

# Témoignage de Stigmatisation ?

## European Survey on Language Competence 2013 – 14 countries



# Niveaux : Common European Framework

The Common European Framework of Reference for Languages (CEFR)

- **B1 and B2: Independent user**

“Importantly, the CEFR levels provide a sound basis for the mutual recognition of language qualifications” (Council of Europe 2008).

Echantillon : 1778 élèves de Year 11 ( âge 16 ans) England ( 55 écoles)  
: 1747 élèves ( âge 16 ans) France (74 écoles)

# Résultats

Table 3.5 Percentage of pupils at each CEFR level for reading in the **first** target language, by jurisdiction

## % CEFR level

Jurisdiction	Pre-A1	A1	A2	B1	B2
England	22.1	57.5	11.2	6.6	2.6
France	28.3	49.0	9.6	7.0	6.1
Poland	27.1	38.1	11.1	10.3	13.4
Portugal	20.2	40.8	12.6	11.1	15.2
Spain	18.0	40.7	11.8	11.8	17.6
Bulgaria	23.1	32.2	11.1	10.2	23.4
Belgium fr	9.7	42.0	17.1	16.5	14.6
Netherlands	3.7	20.8	15.3	22.5	37.7
Malta	3.9	10.2	7.1	15.7	63.1
Sweden	1.4	9.6	8.3	15.1	65.6

% « independent » = B1 + B2

England 9,2%

France 13%

Sweden 80%

# Hypothèses

Il faut changer l'Alphabet Latin (défi sociétal)

Il faut modifier les traditions dépassées  
d'imprimerie sur papier

Il faudra privilégier la lecture à l'écran car à l'écran  
on peut améliorer la lecture avec :

- Annotations en couleur

- Multiplés modifications de taille typographique

- Synchronisations son/texte

- N.B. Émergence de nouvelles technologies de  
synchronisation (SMIL, W3 1999)

# Hypothèses

Il faudra réintroduire l'étude de **la syllabe** dans l'apprentissage de la lecture des L2 car la syllabe est l'élément la plus saillant dans la chaîne acoustique

L'importance accordée aux syllabes au niveau théorique ne trouve pas de traduction sur le terrain de la recherche appliquée

Pour l'enseignement supérieur ceci signifie que la formation des enseignants doit changer et qu'il faut enseigner la lecture des L2 : une activité généralement abandonnée ou négligée par le corps enseignant.

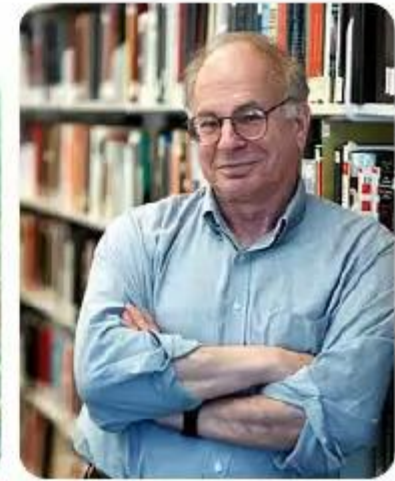
# Hypothèse de base sur un de nos processus mentaux inconscients

L'interférence de la langue maternelle (L1) dans la perception auditive et la production orale de la L2 est la manifestation d'un '**biais cognitif**' (Kahneman)



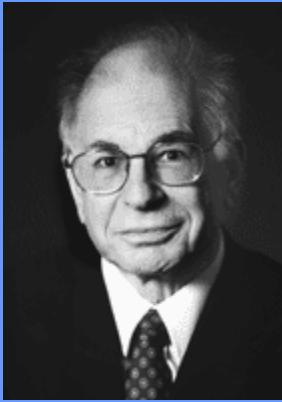
Kahneman &  
Tversky

Biais Cognitif  
1972

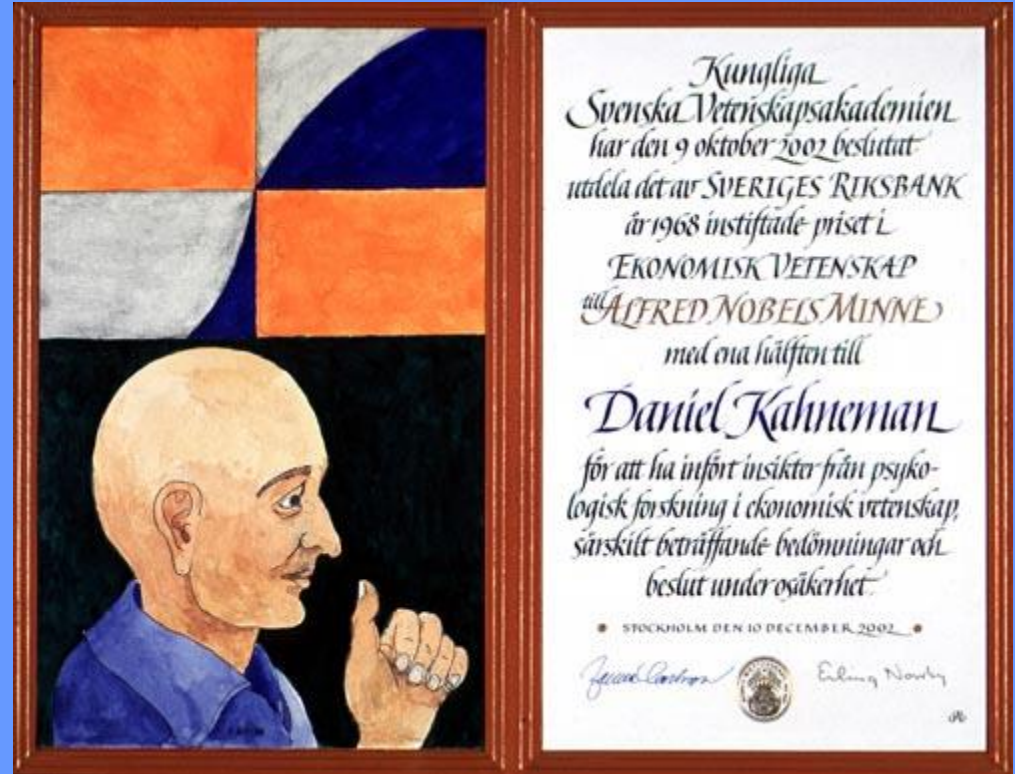


Les **sentiments humains** sont **complexes** et sont fortement **liés au contexte** et aux signaux envoyés par nos perceptions. On distingue la **pensée instinctive** et la **pensée réfléchie**, entre lesquelles il faut trouver le bon rapport. Les **décisions cruciales** doivent faire preuve d'**objectivité en se basant sur des faits** et non des impressions vagues, des influences extérieures ou des raccourcis cognitifs.

# Kahneman & Tversky : le Biaisi Cognitif



•Kahneman & Tversky



•Le Prix Nobel de Kahneman 2002

# Citation de Kahneman

« Notre conviction réconfortante que le monde fait sens repose sur une base solide : notre capacité quasi-infinie à ignorer notre ignorance. »

Plus de la moitié des étudiants d'Havard ont répondu faux à cette question

Le problème :

Une raquette et une balle coûtent 1,10 dollars.

La raquette coûte un dollar de plus que la balle.

Combien coûte la balle ?

Hypothèse :Il faut traiter l'interférence des L1 comme une manifestation d'un biais cognitif

Kahneman donne de l'espoir aux processus de 'Cognitive Bias Therapy' CBT même si les progrès sont lents

Il s'agit d'apprendre à se méfier des prises de décisions hâtives du cerveau rapide.

# Conclusion

Reading AMPLIFIES mother tongue interference in L2 perception and production

# Réactions aux hypothèses

Impossible !

Les intérêts institutionnels liés à cet outil de communication sont trop importants.

C'est inouïe ! C'est trop ambitieux. On ne change pas le conditionnement de toute une jeunesse. C'est trop tard.

# Réponse aux réactions

Il y a un parallèle dans le monde de la  
médecine

La saignée



# Philippe PINEL 1803

La saignée s'arrête en France au 19<sup>ième</sup> siècle après les efforts de Philippe Pinel (1745-1826) (qui avait fait sa thèse à Toulouse et examiné les aliénés enchainés dans l'hôpital la Grave).



# Conclusion

On peut modifier l'apparence de l'AL mais la résistance sera rude, le processus sera lent et dépendra en partie sur le niveau d'interférence générée par chaque L1

# Effets d'annotation/Synchronisation

Sensibilisation

Mémorisation

Perception

Production orale

# Ce qu'on apprend des experts sur la perception humaine

1. La **Psychologie** : l'omniprésence du biais cognitif = nos étudiants peuvent-ils lire la L2 cible ?

Daniel Kahneman : Eugene Higgins Professor of Psychology Emeritus at Princeton University

2. **Neurolinguistique** : notion de « plasticité », le cerveau peut générer des nouvelles structures. Peut-on manipuler le cerveau ?

DEHAENE S. **et al.** 2010. How learning to read changes the cortical networks for vision and language. Science(online). <<http://dec.risc.cnrs.fr/Pdf/col/Dehaene.pdf>>.

3. **Science Cognitive & Art** : La cognition se repose sur les caractéristiques « stéréotypiques », l'art utilise les universaux, le langage utilise les syllables

Vilayanur S. Ramachandran (Director of the Centre for Brain and Cognition, University of California, San Diego)

Plus de la moitié des étudiants d'Havard ont répondu faux à cette question

Le problème :

Une raquette et une balle coûtent 1,10 dollars.

La raquette coûte un dollar de plus que la balle.

Combien coûte la balle ?

# La psychologie sociale

Nous organisons notre perception d'autrui en **simplifiant** les informations qui sont à notre disposition, en les schématisant c'est-à-dire par un processus d'évaluation grâce auquel, au moindre prix nous savons ce que sont les autres et leur mode de fonctionnement.

# Hypothèse

La perception auditive se construit autour d'une **simplification** (perception stéréotypée) de informations disponibles. Cette simplification utilise les éléments les plus saillants dans la chaîne sonore : les syllabes.

# Conclusions 1

Enseigner une L2 sans signaler les différences entre les schémas d'accentuation lexicale de la L1 et de la L2, est une erreur de taille.

Si l'interférence de la L1 est une variante de biais cognitif, la mise en place de CBT (Cognitive Bias Therapy) devient envisageable.



# Evaluation : quelques mesures traditionnelles de pratiques orales et écrites

Etendue générale du champ  
linguistique

Variété du vocabulaire

Maîtrise du vocabulaire

Précision grammaticale

Maîtrise phonologique

Maîtrise de l'orthographe

Justesse au vu du contexte  
socio-linguistique

flexibilité

Prise de  
parole(turntaking?)

Développement  
thématique

cohérence et cohésion

aisance, fluidité à l'oral

précision dans la  
proposition

# Une question fondamentale

Comment mesurer l'intelligibilité des présentations orales des étudiants lorsqu'ils s'expriment dans une seconde langue?

Quelles sont les normes, et comment évoluent-elles (à l'ère de l'anglais universel) alors que l'anglais s'est imposé comme langue internationale, et que les non-natifs sont infiniment plus nombreux que les natifs?

# L'intelligibilité

Bien sûr, c'est une question de bon sens.

Les auditeurs sont habituellement pragmatiques, et tolèrent les forts (ou 'mauvais') accents étrangers.

Lorsqu'un professeur de langue expérimenté ne parvient pas à comprendre, c'est que l'expression orale doit être déroutante (anormale) et « inacceptable »/, irrecevable

# Problème, le contexte de la L1

Comment définir l'interférence de la langue maternelle?

Quels sont les degrés d'“acceptabilité” de l'interférence de la langue maternelle à l'oral?

Comment les seuils de tolérance sont-ils définis?

# « It takes two to tango » Il faut être deux pour danser le tango

La communication est un processus qui se fait dans les deux sens. Les étudiants devraient être autonomes et capables de mesurer eux-mêmes leurs progrès en prononciation. **Mais l'appréciation de l'intelligibilité est rarement démocratique et elle fait rarement l'objet d'une évaluation.** L'analyse de l'enseignement de la langue et de l'évaluation du niveau de langue a occulté un élément important, à savoir la charge inconsciente de préjugés et d'intolérance que l'on retrouve chez de nombreux enseignants et d'évaluateurs à l'égard de certaines formulations orales.

# Les (prochains) défis

Il est essentiel d'analyser (jusqu'à) quel point les évaluateurs du niveau de langue sont influencés par les préjugés dominants, et ignorent les schémas d'accentuation lexicale qu'emploient leurs étudiants. S'il nous faut des descripteurs pour identifier les difficultés rencontrées par les étudiants (ex: "ne peut pas se faire comprendre"), nous avons également besoin de descripteurs pour analyser les difficultés des enseignants (ex: ne perçoit pas le phénomène d'épenthèse et ne peut le tolérer chez ses étudiants d'origine asiatique).

# Le problème

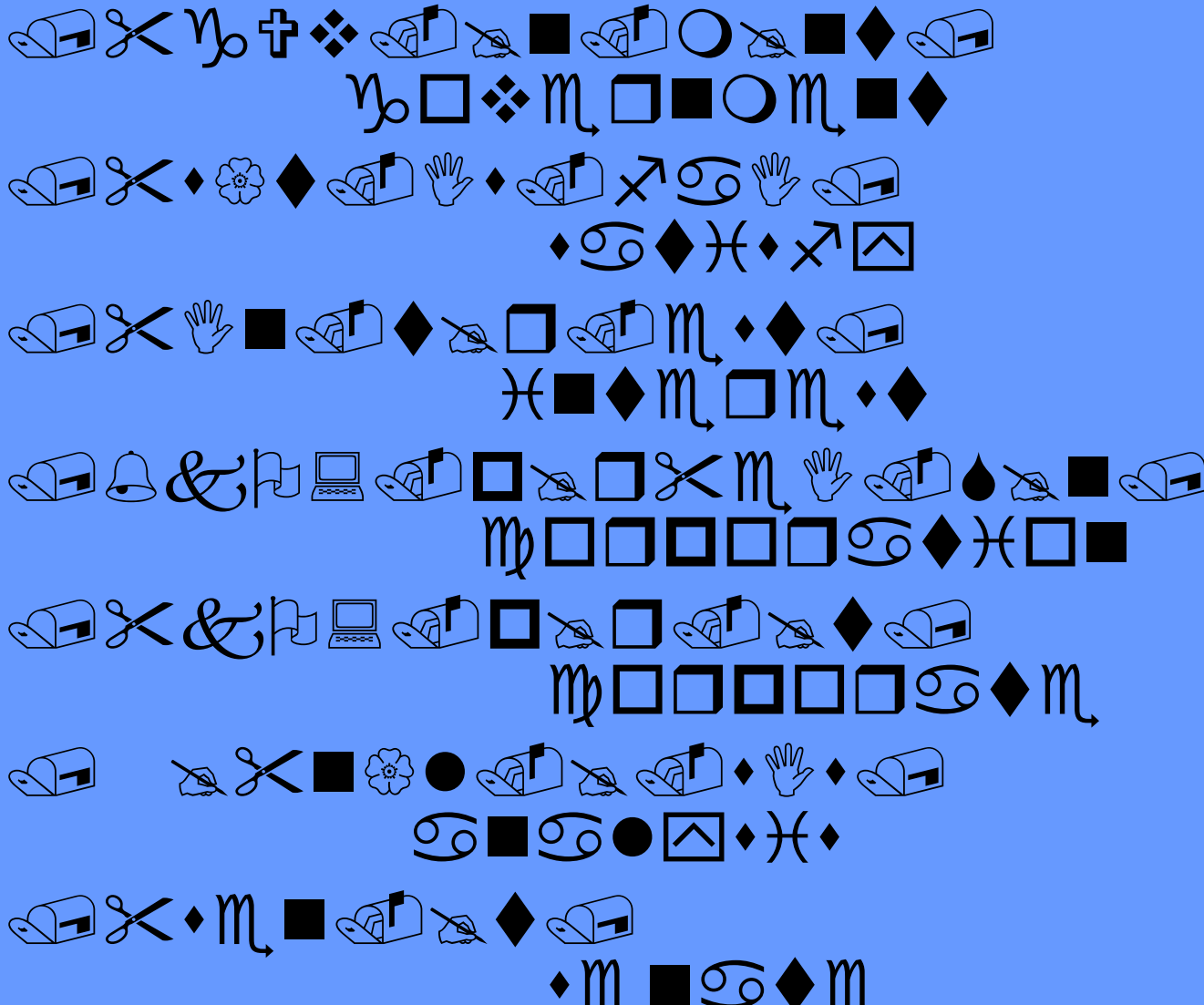
- Les décevantes prestations orales de nombreux étudiants européens.
- Les niveaux d'aisance verbale (la fluidité à l'oral) et les niveaux de précision ne se correspondent pas toujours.
- A la différence de la compréhension auditive, la **perception** auditive, est souvent négligée/occultée. La nécessité de mettre en place un entraînement explicite de l'oreille est souvent évoquée.

# Exemple : Le problème des présentations orales en anglais des étudiants français à un niveau avancé

- Fluidité
- Les étudiants trouvent aisément leurs mots
- Le registre de langue est approprié (correction de la langue)
- Mais...



Sur un total de plus de 250 étudiants de l'université de Toulouse, 70% à 90% se trompent de schémas d'accentuation:



Pourquoi ?

# Pourquoi ?

- Parce que le rôle des syllabes au cours de la lecture en L2 est sous-estimé par les professeurs de langue, ces derniers partant du principe que leurs étudiants savent déjà lire dans la L2 et ont seulement besoin d'exercices sémantiques.
- L'accent est habituellement mis sur la compréhension, et rarement sur les exercices de reconnaissance son/syllabe. Les problèmes de perception liés aux syllabes sont occultés.

# Hypothèse

- L'importance des schémas d'accentuation lexicale en langue étrangère est rarement enseignée de manière explicite.
- Un problème de formation pédagogique des enseignants?
- Un problème de techniques pédagogiques?
- Un problème de métalangue ?

# la syllabe résiste à l'analyse univoque

- La syllabe résiste à l'analyse simple simple et univoque. Elle est souvent décrite comme l'unité la plus saillante et instantanément reconnaissable de la chaîne de sons – un pic central (habituellement une voyelle) entouré de consonnes.

# Le problème de perception à l'oral

- Il n'existe pas de contexte indépendant, ni de propriétés acoustiques immuables susceptibles d'être considérées comme des critères de reconnaissance d'un segment phonologique particulier.

# Le rôle de la syllabe dans la perception

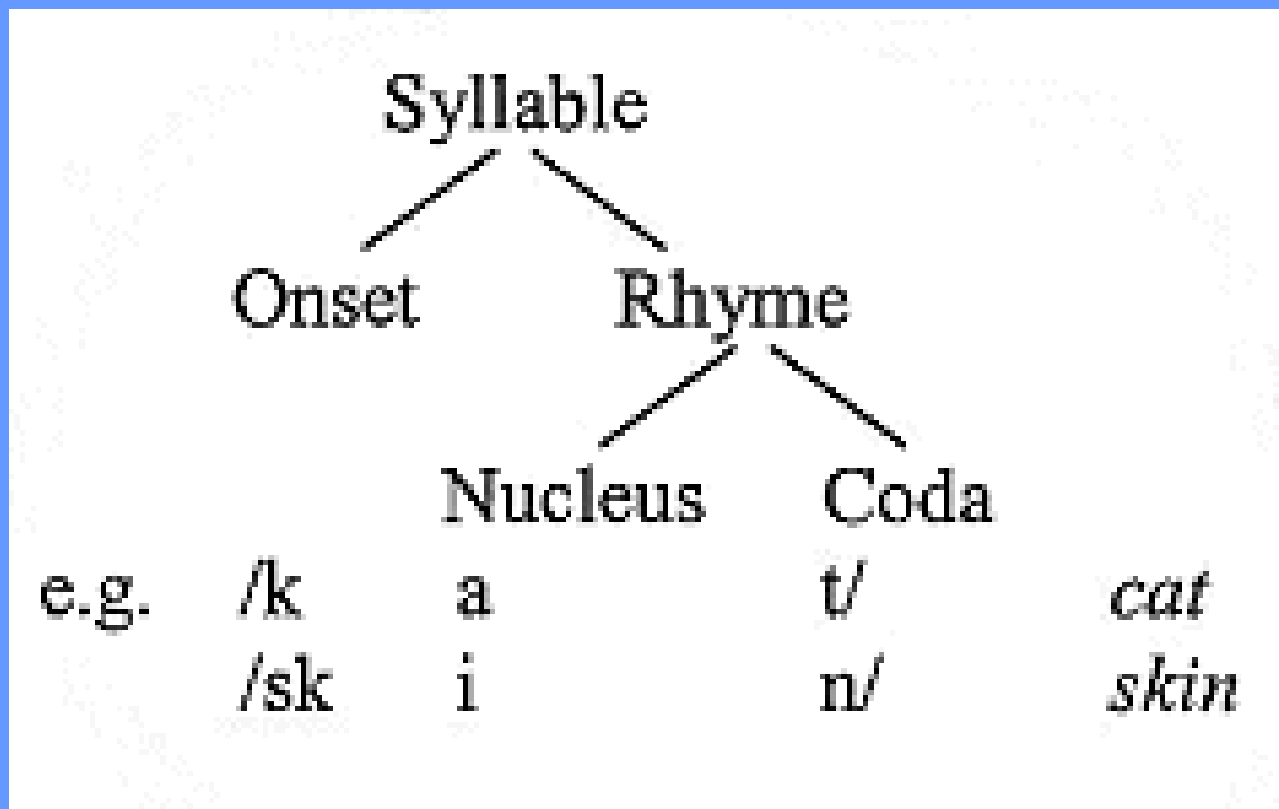
- De récentes études sur la perception ont révélé que les syllabes ont un impact plus grand sur la perception de l'oral que l'on ne le supposait jusqu'alors. (Mattys, S.L. & Melhorn, J.F., 2005).

# L'importance des syllabes

- Chait soutient que les flux de sons de courte durée (~30 ms) et les flux de plus longue durée (~300 ms) sont conçus simultanément et séparément avant d'être réunis dans une représentation stable nommée syllabe. (Chait M. et al 2005).



Cela signifie-t-il que les étudiants devraient se pencher sur les propriétés des syllabes?



Devons-nous connaître les propriétés phonotactiques de la L2? Savoir quels ensembles de consonnes sont possibles, et lesquels ne le sont pas?

- « strength »
- CCCVCCCC

# « Non »

- Les syllabes devraient être étudiées dans le contexte d'un schéma d'accentuation lexicale.
- Comme dans le cas des phonèmes, une bonne dose de reconnaissance intuitive mènera loin. Il importe surtout d'avoir quelques connaissances de base concernant les principaux problèmes d'interférence linguistique, à l'instar des difficultés rencontrées par les étudiants français lorsqu'ils s'expriment en Anglais.

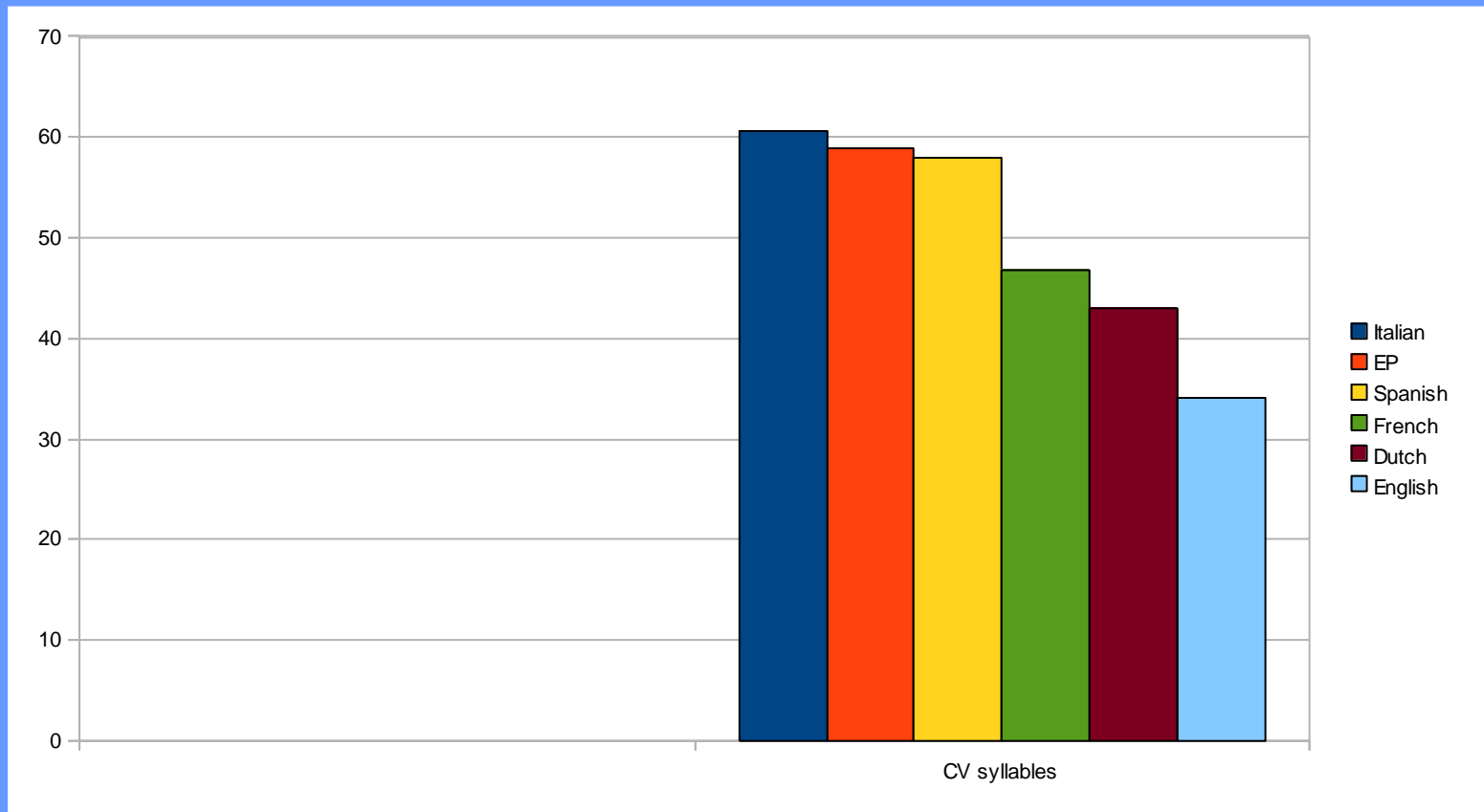
# Quelques notions concernant la complexité syllabique

- Le nombre des différentes syllabes permises et les façons dont elles peuvent être disposées
- Globalement, quel est le degré de difficulté pour déterminer les limites des syllabes dans une langue donnée?
- Fenk(2008) définit la complexité syllabique comme le nombre de phonèmes par syllabe.

# La complexité syllabique

- Cutler (1986) soutient que la langue française possède des mots que l'on peut aisément décomposer en syllabes, ce qui ne serait pas le cas en Anglais.

# CV syllables (Adsett & Marchand, 2010)



# Facultés de lecture des mots inventés (non-words)

- Les taux d'erreur sont significativement plus bas chez les enfants (de niveau CP et CE1) dont la langue maternelle est le finlandais, le français, le grec, l'italien, l'espagnol, et le portugais. (Toutes sont considérées comme des langues ayant une structure syllabique simple.)
- Les taux d'erreur sont sensiblement plus élevés concernant les Autrichiens, les Danois, les Hollandais, les Anglais, les Allemands, les Islandais, les Norvégiens, et les Suédois. (Toutes ces langues sont considérées comme ayant une structure syllabique plus complexe.)  
Seymour et al. 2003

# La prononciation des voyelles en anglais est imprévisible

Par exemple , “a”

se prononce de façon différente dans

“*cat*”, “*call*”, “*car*”, “*cake*”, and “*care*”.

La séquence « ough » est également complexe :

/oʊ/ as in "though" (cf. *toe*), /u:/ as in "through" (cf. *true*).

/ʌf/ as in "rough" (cf. *ruffian*), /ɒf/ as in "cough" (cf. *coffin*), /ɔ:/ as in "thought" (cf. *taut*),

/aʊ/ as in "bough" (cf. *cow*): /ʌp/ as in hiccough ( cf. “cup”), /ə/ as in "thorough" ( cf. *borough*).



# L'interférence de la langue maternelle (1)

- L'interférence de la L1 inhibe la perception des schémas d'accentuation de la L2, ainsi que la perception des syllabes fortes et faibles.
- L'accentuation française est régulière et se fait sur la droite. On parle dans ce cas d'une langue « codachrone ».
- L'accentuation anglaise est irrégulière et se fait sur la gauche. On parle d'une langue « capochrone ».

Les problèmes d'interférence soulevés par les schémas d'accentuation lexicale anglais sont-ils les mêmes pour toutes les langues?

- Bien sûr que non
- En tchèque, letton, hongrois, allemand suisse, (dialecte Bernois), finlandais et swahili, l'accentuation se situe toujours sur la première syllabe. (Le finlandais n'a pas de réduction de voyelle non accentuée.)
- En français, espagnol, portugais, turc et polonais, l'accentuation se fait rarement sur la première syllabe.

# L'interférence de la langue maternelle(2)

La qualité de l'accentuation varie selon les langues:

La hauteur (pitch)

La longueur

Le volume

Un phénomène complexe qui comporte plusieurs paramètres.

# Un phénomène universel

- L'interférence de la langue maternelle dans la perception et la production est un phénomène universel que l'on nomme «engagement neuronal », mais il est **amplifié** par l'alphabet LATIN, qui utilise les mêmes symboles pour décrire des systèmes de sons différents.

## Le problème de l'alphabet latin (3)

- La lecture et la sous-vocalisation (la lecture silencieuse) peuvent aggraver/renforcer les mauvaises habitudes. L'alphabet latin est un outil plurilingue développé et adapté en fonction de chaque langue par des mentalités **monolingues** et ethnocentriques.
- L'interférence est intimement liée à la proximité morphologique. Plus deux langues sont proches et plus l'interférence sera grande en raison du recours aux mêmes symboles latins pour décrire différents schémas d'accentuation lexicale.

# Le problème de l'alphabet latin

- Ne pas tenir compte des schémas d'accentuation lexicale revient à négliger/occulter les éléments acoustiques les plus essentiels de la chaîne de sons.
- L'accentuation s'accompagne d'une plus grande amplitude respiratoire.
- Hypothèse : Apprendre une L2 sans consciemment analyser le rôle des schémas d'accentuation lexicale pourrait tout simplement renforcer des problèmes d'interférence, ainsi «fossilisés » **à vie.**

# L'interférence de la langue maternelle

- Conclusion
- L'anglais et le français sont des langues morphologiquement amies, mais acoustiquement ennemies.
- 60% de la langue anglaise est composée de français ou de latin.
- Les schémas d'accentuation lexicale sont radicalement différents.

# Conséquences de l'interférence de la langue maternelle

- Rupture et échec de la communication
- Intolérance = les préjugés sont confortés et renforcés
- Stigmatisation
- 'Echec' comparatif (la France est 69ème sur 120 nations dans un classement mondial suite aux tests TOEFL de 2009)



# Conséquences de l'interférence de la langue maternelle (2)

- Les Anglais et les Français sont mutuellement démoralisés par leur incapacité respective à prononcer la L2 correctement. Ils sont victimes de l'alphabet latin, mais aussi d'une intolérance généralisée à l'égard de l'interférence de la langue maternelle.
- Au lieu de remettre en cause l'alphabet latin, ils se tiennent habituellement eux-mêmes pour responsables, invoquant des facteurs culturels (« perfide Albion ») ou blâmant un enseignement traditionnel jugé incompetent du fait de son parti pris littéraire.

# Conclusion

- La lecture AMPLIFIE l'interférence de la langue maternelle dans la perception et la production de la L2.

« Engagement neuronal »+ lecture  
= majoration de l'interférence

- Il faut des années au cerveau pour acquérir des facultés de lecture rapides, automatiques et sans erreur.  
L'apprentissage en L2 de différents codes associant les sons pour décrire les mêmes symboles induit une résistance, et majore de ce fait l'interférence.

# Conclusion

- De nouvelles méthodes sont nécessaires pour visualiser le son, et libérer ainsi le lecteur des confusions entre son et symbole liées à l'utilisation du MEME alphabet latin pour de nombreuses langues.

SWANS

Synchronised Web  
Authoring Notation  
System

# Fields of study of SWANS project specialists

- Linguistics
- Auditory Perception
- Cognitive Psychology
- Computer Science
- Document design specialists

# Laboratories involved

- Laboratoire Interuniversitaire de Recherche en Didactique des Langues, (LAIRDIL) Toulouse,
- Anthony Stenton, Anne Péchou, Gail Taillefer, Nicole Décuré, Antoine Toma, Christine Vaillant Sirdey

- Laboratoire Travail et Cognition  
UMR CNRS 5551, Toulouse
- André Tricot
  
- **Laboratoire de Neuropsycholinguistique**  
**Jacques-Lordat**, EA 1941 Toulouse
- Pascal Gaillard, Michel Billières
  
- Université Toulouse 1 : Aryel Beck



- **Laboratoire d'Analyse et d'Architecture des Systèmes, LAAS CNRS 7, Toulouse**
- Saïd Tazi

**Equipe Automatique des Procédés Biotechnologiques, Laboratoire de Biotechnologies-Bioprocédés, INSA Toulouse**

- Nabil Kabbaj

# Project objectives

- Measurable progress in lexical stress perception and oral production
- The development of an authoring system for synchronising and annotating texts for an adaptive multilingual context
- Allowing teachers to annotate without having to master advanced programming skills or spend hours studying stress patterns= semi-automatic or automatic annotation

# The challenges

- To define the rules of English stress (Alain Deschamps 94, d'après Guierre) and an appropriate learning method (Théo van Leeuwen : Visual Design, Rudolph Arnheim : Visual Weight)*
- To develop protocols for evaluating learning (Sweller, Mayer, Paivio)*
- To develop an 'Intelligent Environment for Human Learning' where visual perception can help improve auditory perception*
- To develop a prototype authoring system to make the process of textual annotation semi-automatic*

# Hypothesis of dual coding (Paivio)

- Dual coding (sound + text + paralinguistic markers) should improve learning for novice learners while having no effect or a negative effect for those who already know the place of the accent.

# Textual annotation techniques for learning pronunciation

- How far can we go ?
- When does the brain start to complain because reading becomes too slow or because codes seem too obscure ?
- What constitutes cognitive overload ?

# Can typography change ?

- Is our software capable of offering new more memorable techniques for learning pronunciation ?
- Are we ready for a more flexible, animated and colourful alphabet ?

# A proposed solution

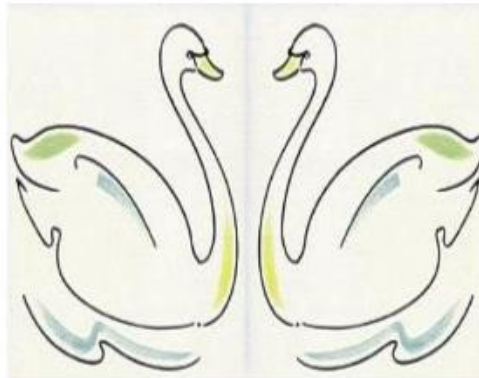
An authoring system for improving pronunciation

SWANS Configuration

Import Word (HTML file)

Generate Web Page(s)

## SWANS



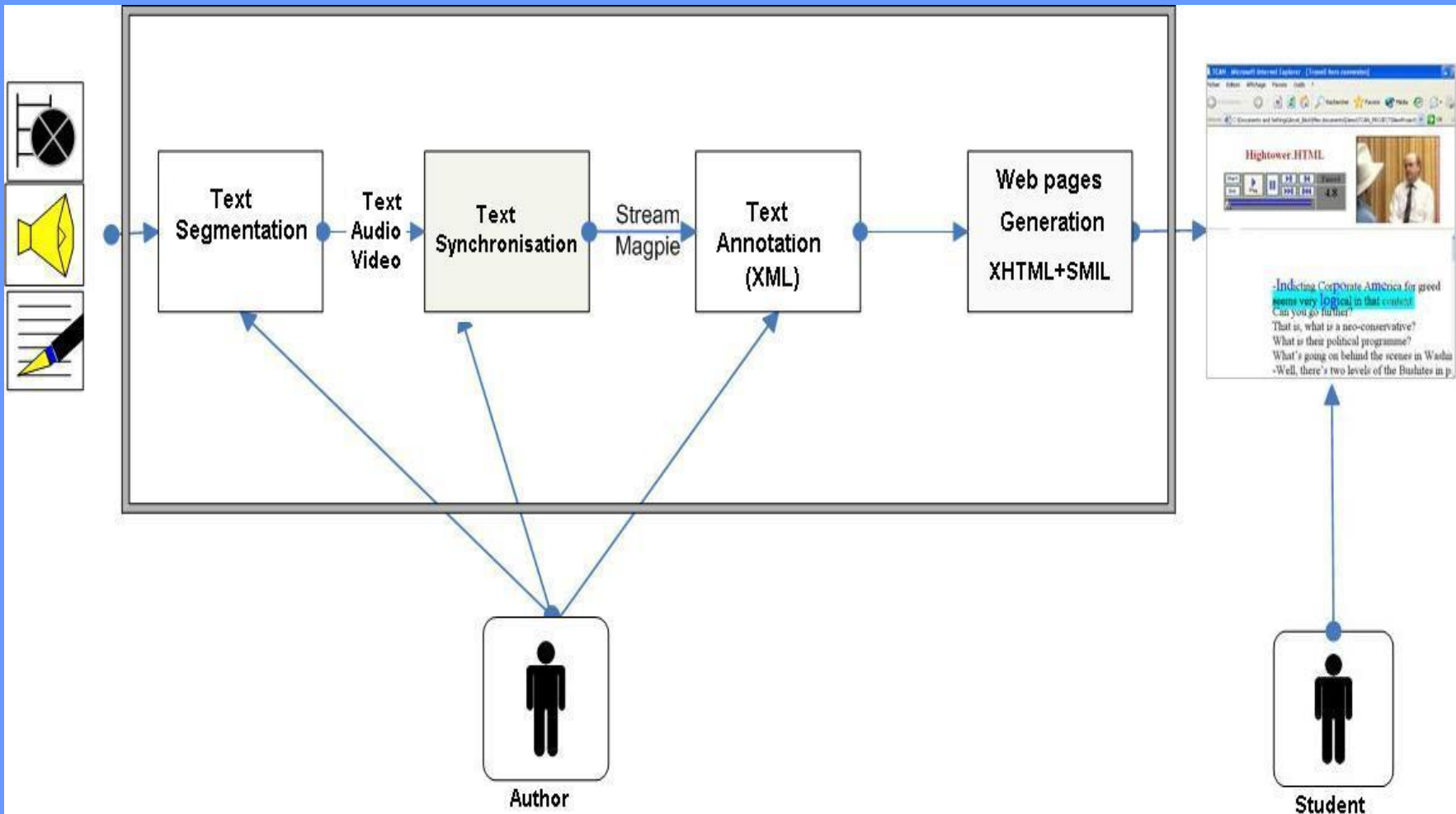
Synchronised Web Authoring Notation System

# Annotated text in Word before web page generation in SWANS

YouTube is one of the Internet's fastest success stories  
and it's two OWners now stand to join the ranks of the world's super rich  
after Google agreed to buy it for almost £900 million.  
“We just want to say thank you.  
Today we have some exciting news for you.  
We have been acquired by Google !”  
The social networking site was created 18 months ago  
by a pair of ‘20-something’ Californians.  
It has just 67 staff and is based above a pizza restaurant.  
But it's become a worldwide phenomenon.  
So what makes YouTube such a hit ?  
Well, it's all about turning orinary people into celebrities|



# The four stages of authoring with SWANS



# Annotation techniques

Hem, there is a **C**ertain,  
**d**efinitely  
a **C**ertain  
re**a**lity to that.  
There is a very ...  
di**St**inct **d**ifference  
be**t**ween what  
**A**merican **b**usinesses,  
and ... **A**mericans in general,  
ex**p**ect from their  
**g**overnment,  
and what French **b**usinesses

# A text annotated with Swans

**Read and Listen**  
Marcel Berlin, Law in action  
on LEGAL AID. Nov 04

Text Accent

His **p**essimism, though is shared by **l**awyers **a**cross the **c**ountry in the **L**egal Aid field.

**L**ord **P**hilip of **S**idbury, a long standing **s**olicitor and **L**iberal **D**emocrat peer has been **a**mong the most **v**ocal **c**ritics of **g**overnment **p**olicy.

It's all **p**retty sad, we did once **b**oast the **f**inest **L**egal Aid **s**ystem in the world, one of the great **l**andmarks of the **p**ostwar **A**tlee **g**overnment, and to be **h**onest,

Start End Play Paused 6.7

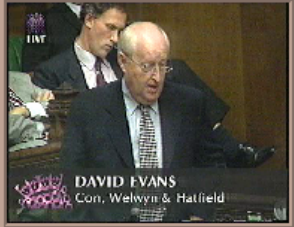
BACK NEXT MENU EXIT

# Question Time in Parliament annotated with SWANS

Question Time in Parliament - Microsoft Internet Explorer

This page

Read and listen  
Question Time in Parliament



DAVID EVANS  
Con. Welwyn & Hatfield

Text Accent Intonation Karaoke  ON  OFF

Mr David Evans (Conservative M.P.)  
Will...Will my Right Honorable Friend remind me and all those people under thirty five years of age..... Was it a Conservative [1] Government that let inflation rip to twenty six point nine percent? Was it a Conservative Government that allowed a higher rate of tax to go to ninety eight pence in the pound? Was it a Conservative Government that did not pay the Senior Citizens Christmas Bonus? And was it a Conservative Government that...

Start End Play Paused 7.2

BACK NEXT MENU EXIT

# A Spanish text

## Cumbre Iberoamericana en Bolivia3.html



Text Accent

X.R.: En América Latina, a día de hoy, el 10 % de la población acapara la mitad de la riqueza, y según el Banco Mundial, cada vez es mayor la brecha que separa a los ricos de los pobres en esta región del mundo. De cómo luchar contra estas desigualdades es lo que ha tratado la XIII Cumbre Iberoamericana celebrada en Santa Cruz de la Sierra en Bolivia el pasado mes de Noviembre. Allí se reunieron más de 5000 personas entre diplomáticos, jefes de estado de 21 países, técnicos y periodistas. Ha habido ausencias notables como la de Fidel Castro, por tercera vez consecutiva.

# An annotated German text

## Clara Schumann



Text Accent



reich die **Darbiere** **deci**, die **gech** werden.

Clara Schumann ist den **meisten** nur durch den **Hundert-Mark-Schein** vertraut.

Sie war eine **große** Komponistin im **Schatten** ihres **Mannes**.

Das Bonner **Frauenmuseum** veranstaltet **ihr** zu Ehren eine **Ausstellung**.

**Clara Wieck** als gefeierter **Kinderstar** im **Alter** von **dreizehn** und **fünfzehn** Jahren.

Als **heimliche** Verlobte von Robert **Schumann** mit **neunzehn** Jahren.

**Fünfunddreißig** Jahre alt, sie hat nun für **sechs** **Kinder** zu **sorgen**.

Robert, ihr **Ehemann**, ist oft **krank** und hat **Depressionen**.

Das **Haushaltsgeld** ist **knapp**. In **dieser** **Zeit** wird ihr Klavierspiel zur **wichtigsten** **Einahmequelle** für die **Familie**,



Cimabue 1300



Rogier van der Weyden 1450

Does reading on a computer screen or nomadic devices (i-pads,smart phones) change brain organisation ?





# How does memory behave after dual coding ?

Département des Langues et Civilisations, Université Toulouse I

"Lies, Damned Lies and Statistics" : BBC



*Oh yes, the **opinion** polls, you know the ones the **politicians** always **ignore**, the ones they never pay **any attention** to. **Ever**. And yet they always seem rather well **versed** on the last **de**tail. It's funny that. But the **opinion** poll is an **extre**mely valuable tool, right for misinterpre**ta**tion, acci**den**tal or otherwise... The pollsters **strive** to keep their **sci**ence as **reli**able*

# Should synchronised reading be intensive or extensive ?

**BBC2.HTML**

Start End Play Paused 13.0

Text Accent

Permet de visualiser les accents ou de voir le texte "nu"

Permet de suivre le son

Politicians may be worried and businesses around the world unhappy but foreign tourists in New York can barely contain their delight at the weak dollar. It means they get to stay in posh hotels, eat in the best restaurants and shop to their heart's content, all without huge credit card bills when they get back home. The BBC's Emma Simpson has been hitting the tourist trail. New York at Christmas there's nothing quite like it. This is high season from Manhattan. So this month a night at the Waldorf will cost you around \$600. Pricey, but here they're always busy. This year, though, foreign visitors are up by a third.

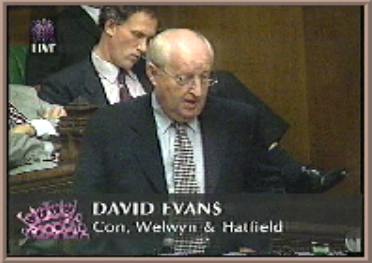
Terminé Poste de travail

# Does synchronised reading inhibit 'L1 interference' during silent reading?

Question Time in Parliament - Microsoft Internet Explorer

This page

Read and listen  
Question Time in Parliament



DAVID EVANS  
Con, Welwyn & Hatfield

Text Accent Intonation Karaoke  ON  OFF

Mr David Evans (Conservative M.P.)  
Will....Will my Right Honorable Friend remind me and all those people under thirty five years of age..... Was it a Conservative [1] Government that let inflation rip to twenty six point nine percent? Was it a Conservative Government that allowed a higher rate of tax to go to ninety eight pence in the pound? Was it a Conservative Government that did not pay the Senior Citizens Christmas Bonus? And was it a Conservative

Start End Play Paused 7.2

BACK NEXT MENU EXIT



# How does memory behave when text moves ?

Three « levels of representation for words in memory » (= triple storage):

- conceptual level
- semantic-syntactical level (*lemmes*)
- formal level (*lexèmes*)
- Willem Levelt, **Modèle de la production langagière**
- (résumé from de Bot, Paribakht & Wesche 1997, 312)

Levelt's analysis is based on student reading of **paper** documents

- Toulouse tests of dual coding at the word level and at the sentence level constitute an early and limited analysis of the potential of the new plasticity in computer based reading
- Over 200 students participated
- Sound recordings of student pronunciation were stored to examine stress perception

Problem : The adult brain is conditioned by typography

- « *Freeing ourselves from chirographic and typographic bias is probably more difficult than any of us can imagine* » Ong 1982

# Feedback from field tests

- Feedback concerns the pedagogical use of documents created with SWANS in French and European centres
- Feedback on the technical aspects of the automatic authoring system is due in 2016



# Pedagogical approach : Technique 1

- *The student is invited to **annotate** the scripts on paper to indicate:*
- Primary accents (and secondary accents)
- Weak forms

# Examples

- **Student answer**

- **Solar Energy**

- Atmos**ph**ere

- Elect**ricity**

- **Bioclimatic**

- Env**ir**onment

- Deter**mi**ne

- **Suggested right answer**

**Sol**<sub>a</sub>**r** **En**<sub>e</sub>**r**gy

**Atm**<sub>o</sub>**s**phere

Elect**ricity**

**Bioclimatic**

Env**ir**<sub>o</sub>**n**<sub>e</sub>**m**ent

Deter**mi**ne

# Technique 2 : Reading Aloud

- *Pair work*
- *One student reads out loud from an unannotated text*
- *The other listens and corrects the mistakes while reading the annotated text*

# Results

- Correct keyword annotation does not always mean correct oral performance
- Wrongly annotated keywords do not always mean errors in oral performance

# Results

- The approach gives positive results for most presentations : improved intelligibility, listener's temptation to 'switch off' diminishes
- Student motivation increases as does the demand to learn stress rules
- Learning is relatively unstable. Random stress patterns return when difficulty or tension increases

# Technique 3

- The Carrousel approach finds a new justification
- Correct oral performance in controlled conditions is no guarantee that the correct stress patterns will be used in spontaneous speech

# Authoring Performance

- The prototype SWANS was used for the generation of several hundred synchronised and annotated documents (English, Spanish and German) 2005 to 2007.

# Performance

- The creation of synchronised annotated documents can now be carried out by non specialists but still requires time we can ill-afford to spare
- Creating an **automatic stress pattern annotating tool** is a major challenge akin to the automatic spelling correction programmes developed over 30 years between 1965 and 1995.



# Conclusions

- Synchronised reading based on dual coding techniques opens promising new perspectives for improving perception and pronunciation.
- The creation of several tailor-made documents using the same basic audio or video with text annotations adapted to particular kinds of L1 interference presents a new field for collaborative on-line networked teacher training

# Can annotation become automatic?

- Like spelling correctors accuracy levels are variable but 'acceptable'
- Method 1 rule-based expert system
- According to Deschamps we can deduce the stress pattern of a word from its spelling in 90% of cases. Some rules, such as those concerning endings with <\_ic, itis, osis> or the <ION> rule, are close to a 100% (Deschamps p.33, 1994).

# Can annotation be automatic ?

- **Method 2 voice recognition techniques**

The integration of voice recognition techniques giving raw and optimised phonemic transcriptions adapted from the PERL algorithms of the Aix-Marsec project is a more complex challenge. The Aix-MARSEC project time-aligns speech data at the phoneme and syllable level. Aix- MARSEC tools consist of a set of reference files (grapheme-phoneme conversion dictionaries) and multiplatform Praat and Perl scripts.

# Can annotation become automatic?

- Method 3 text banks of annotated texts
- a bank of pre-annotated scripts, from 2004-2005 3-minute video news bulletins, offers an alternative to rule-based or dictionary based systems (e.g. 71,000 entries of the Oxford Advanced Learner's using the SAMPA alphabet) but, of course, requires manual adjustment for ambiguous cases ("*convict*" = the verb /&✂■✂❖✎&◆/ or the noun ✎✂&➔■✎❖✎&◆/ ?) .

# Automatic Syllable annotation

- A syllabifier is an algorithm for the automatic detection of syllables
- Problem : **no accepted standard algorithm for automatic syllabification exists.**
- According to the circumstances (such as variable rate of enunciation), a given phoneme sequence may be syllabified differently.
- **Rule-based and statistical approaches** are common (see Hammond's and Fisher's implementation of Kahn's original research in his well-known 1976 thesis) as are **data-driven approaches** based on machine readable dictionaries or databases of known syllabifications.

# But....

- Machines can syllabify as well as experts and do so more quickly!
- Highly successful syllabification can now be performed either on the **orthographic representation of words** or on the **phonetic transcription** of words, according to the particular task required.

# Maximum onset principle

- A common rule, that is often cited, is that more consonants are generally present in the onset of a syllable than in its coda; this is known as the *maximal onset principle*. Such knowledge can be built into a syllabifier.

# Existing machine-based research (1)

- Schmid presents a statistical approach to syllabification in **German** and claims an accuracy level by word of **99.85%** (Schmid et al, 2007).
- Adsett & Marchand's syllabification by analogy procedure achieved **97.70 %** accuracy for **Italian** whereas the best rule-based methods attained **89.77%.**
- In the United States, the ToBI framework (Ohio State university) has been used for transcribing the intonation and prosodic structure of spoken utterances in several languages notably **English and German.**



# Existing machine-based research (2)

- For the **French** language, Avanzi et al. have used the Dutch programme PRAAT to achieve **an effective degree of syllable recognition accuracy** (Avanzi et al. 2007).
- Marchand and Damper have developed an algorithm for determining syllable boundaries in the orthographic form of unknown words that works by analogical reasoning from a corpus of known syllabifications (Marchand & Damper, 2009).
- **Data-driven techniques outperform the rule-based systems** in word and juncture accuracies by a very significant margin in the case of English but require training data and are slower.

Can the TTS expertise of machines be extracted to serve European humans ?

- Like the development of a reliable on-line language tests for all European languages (a more reliable DIALANG ) syllabifiers must now be considered a Europe-wide research priority
- It is not a question of making machines more intelligent but of making humans more knowledgeable.

# TTS algorithms

- Syllable structure information is used in text-to-speech (TTS) systems for Czech, European Portuguese, German, Italian and Romanian among others (Adsett, 2009 ).

# The 'extensive reading hypothesis'

- Extensive reading of synchronised text has helped students to inhibit interference from the L1 in controlled contexts but not in spontaneous conversation
- Perception of stress improves in novice learners exposed to sound and annotated text.
- Feedback from students suggests explicit knowledge of stress patterns is a new **springboard** for turning *savoir* into *savoir faire* and *savoir être*.

# Conclusions

- The creation of reasonably accurate syllabifiers for all European languages is a research and pedagogical priority. If the budget for this research exceeds the traditional 400,000 euros Socrates ceiling we need to change the ceiling not the research.
- SWANS offers a prototype authoring system with implications for learning and teacher training. Feedback suggests it deserves wider testing at novice and intermediate levels.

# Does SWANS announce the Swan-Song of the Gutenberg tradition for language learners studying pronunciation ?

- Resistance to changing reading habits may ensure an erratic evolution of reform but we believe the way is open for new experiments with intensive and extensive reading on the computer screen with potentially important repercussions for perception and oral production.

# Will teachers use SWANS ?

- The eternal debate ... should teachers become authors of on-line resources ?
- Testing is taking place in several European language centres affiliated to CERCLES. Feedback suggests “yes” if annotation becomes automatic and as simple as present day spelling correction.

# Invitation

- To participate in a European project to introduce automatic annotation of lexical stress patterns or for further information on SWANS please contact :

[anthony.stenton@univ-tlse1.fr](mailto:anthony.stenton@univ-tlse1.fr)



# Que faire?

## Éléments pour une discussion

Sensibilisation :

Réintroduire l'étude de la syllabe dans le contexte de l'apprentissage des langues étrangères

Valoriser l'étude des schémas accentuels des langues nationales européennes

Mesurer les progrès des étudiants dans la perception et la production orale des L2

Introduire une formation sur la perception auditive et MTI pour les enseignants qui sera plurilingue (langues européennes et les langues importantes dans le monde).

Mettre en place un réseau de partage de documents annotés selon la nature des interférences de chaque L1 de chaque étudiant

Soutenir financièrement les universités qui cherchent à développer des algorithmes d'annotation automatique pour toutes les langues européennes

Faire du plurilinguisme un atout

Sensibiliser les établissements d'enseignement supérieur à l'intérêt de traiter les questions de langues dans le cadre de politiques linguistiques

Élaborer un *Guide des bonnes pratiques linguistiques dans les universités*