COMPUTATIONAL PSYCHOLINGUISTICS AND SPOKEN WORD RECOGNITION IN THE BILINGUAL AND THE MONOLINGUAL

Thèse présentée à la Faculté des Lettres et Sciences humaines
Institut des Sciences du langage et de la communication
Université de Neuchâtel

Pour l'obtention du grade de Docteur ès Lettres
Par

NICOLAS LÉWY

Acceptée sur proposition du jury:
Prof. LOUIS DE SAUSSURE (Université de Neuchâtel), directeur de thèse
Prof. GARETH GASKELL (University of York, Angleterre)
Prof. ELSA SPINELLI (Université de Grenoble, France)
Prof. MICHAEL THOMAS (Birkbeck, University of London)

Soutenue le 26 juin 2015

Université de Neuchâtel
2015

UNIVERSITÉ DE NEUCHÂTEL

Faculté des Lettres et Sciences Humaines
Espace Louis-Agassiz 1
CH - 2000 Neuchâtel
Tél : +41 (0)32 718 17 00
Fax : +41 (0)32 718 17 71
secretariat.lettres@unine.ch

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de M. Louis de Saussure, directeur de thèse, professeur à l'Université de Neuchâtel ; Mme Elsa Spinelli, professeure, Université de Grenoble ; M. Gareth Gaskell, professeur, Université d'York ; M. Michael Thomas, professeur, Birkbeck College, London, autorise l'impression de la thèse présentée par M. Nicolas Léwy en laissant à l'auteur la responsabilité des opinions énoncées.

Neuchâtel, le 26 juin 2015

La doyenne
Geneviève de Weck

par ordre

Keywords:

computer simulation, speech perception, word recognition, French language, bilingualism, lexical access, psycholinguistic models.

Mots-clés:

simulation sur ordinateur, perception de la parole, reconnaissance des mots, langue française, bilinguisme, accès au lexique, modèles psycholinguistiques.

ABSTRACT

This is a doctoral thesis in computational psycholinguistics, an interdisciplinary research field combining expertise and experiences in linguistics, psychology, and computer science. The thesis takes as a subject the simulation of human word recognition, that is, it aims at modeling on the computer the cognitive process of how we activate and access words, their forms, and their constituting components, in our mental lexicon. It concerns spoken (rather than written) word units and consequently deals with the simulation of auditory (in contrast to visual) word recognition. Specifically, this thesis puts forward two new models that we have developed and built ourselves: One is named FN5 and simulates spoken word recognition in monolinguals; the other, called BIMOLA, models spoken word recognition in bilinguals. The monolingual model FN5 is on French and therefore contains a lexicon of 17,668 French words (nouns, determiners, and prenominal adjectives), some of which having variants and giving rise to a total of 20,523 pronunciations. FN5 processes single (i.e. isolated) words as well as sequences of two connected words (determiner + noun, or prenominal adjective + noun). It implements a new approach to recognizing sequences of words, by means of optimizing the words' alignment positions and pronunciation variants. In addition, it provides for several phonological phenomena that can take place within a word or at boundaries between words (deletion of schwa, linking with and without liaison, word contractions including elision). To account for dialectal differences, it may be run in either of two versions, standard French or Swiss French. The bilingual model BIMOLA deals with two languages all at once, English and French; it includes an English–French bilingual lexicon of 8,696 words (all verbs, 4,348 for each language); and it operates in various language modes (i.e. global configurations of the bilingual's two languages). BIMOLA is able to identify words from either language (always single words, to keep things easy), including guest words, that is, code-switches and borrowings from one into the other language. The two models share a phonetic feature matrix that represents similarities and differences of phonemes both within and between the languages and dialects. As revealed in the evaluations, both our models have a great overall recognition performance and are able to simulate a large number of specific psycholinguistic effects.

ACKNOWLEDGMENTS AND DEDICATION

I thank Prof. François Grosjean, my original thesis advisor, for the stimulating, innovative, and demanding work at the Laboratoire de traitement du langage et de la parole (Language and Speech Processing Lab), at which I was lucky and very happy to participate first in a research project in computational linguistics (funded by CERS/KWF) and subsequently in a total of four research projects in psycholinguistics (all supported by the Swiss National Science Foundation). In the SNSF projects no. 1214-058848 (2000 to 2003) and no. 100012-103384 (2004 to 2006), entitled “Elaboration d’un modèle psycholinguistique de la reconnaissance des mots dans la chaîne parlée”, an earlier version of the FN5 model was conceived and developed by myself (under guidance of F. Grosjean). I had the pleasure to team up with my colleagues, Lysiane Grosjean, Isabelle Racine, and Carole Yersin, who checked and corrected the whole lexicon, translated the interface into French, and did numerous informal evaluations (not reported here); Sandra Schwab helped me find stimuli for a first systematic evaluation (the evaluations and statistical analyses presented in this thesis were, of course, all conducted by myself, from 2007 to 2014, and are my sole responsibility). I also would like to thank Jacqueline Gremaud-Brandhorst, a British English native, for kindly examining BIMOLA’s English lexicon.

After many years of supervising my doctoral thesis, François finally was obliged to withdraw from it for health reasons, which I very much understand. It left me, unfortunately at this stage, with a near-to-ready dissertation and no supervisor. I am therefore very grateful to Prof. Louis de Saussure for his readiness to step in as my final thesis advisor, even if my research topic is outside his usual focus of interest, and in doing so, to allow me to bring my dissertation to completion. Interacting with him, during this last phase, was efficient and very pleasant. My sincere thanks also go to Prof. Gareth Gaskell, Prof. Elsa Spinelli, and Prof. Michael Thomas, for generously and promptly agreeing to participate in my jury and for providing some excellent suggestions on how to amend my manuscript.

I dedicate this to a dear friend and great artist, who so far holds no fewer than 17 honorary doctorates, and to my mother who needs none to be my best friend.

OUTLINE

Chapter 1, “Introduction”, provides a historical survey of the research topic of spoken word recognition modeling and reviews important models proposed by previous researchers in computational psycholinguistics.

Chapter 2, “Two new models: FN5 and BIMOLA”, gives the reasons why the two models that we have elaborated ourselves are desirable and presents their general architecture. It makes clear which elements of this thesis are relevant to both models and which are specific to one or the other model.

Chapter 3, “Linguistic knowledge”, identifies the linguistic information that the models possess at the level of features, phonemes, and words. It introduces the feature matrix, the phoneme repertories and metric space of phonemes, and the lexicons used by BIMOLA and FN5, respectively.

Chapter 4, “General mechanisms”, presents the internal mechanisms that are common to both models: the activation and inhibition of phonemes, as well as the activation, inhibition, and isolation of words.

Chapter 5, “Specific mechanisms”, describes the sequential processing mechanisms that belong uniquely to FN5 and the language activation mechanisms that are pertinent to BIMOLA.

Chapter 6, “Evaluating FN5”, explains the evaluation method and tools and reports on our evaluation of FN5. There are simulations on isolated, single words and simulations on sequences of connected words.

Chapter 7, “Evaluating BIMOLA”, gives account of our evaluation of BIMOLA. There are monolingual simulations (English or French) and bilingual simulations (dealing with both languages together).

Chapter 8, “Conclusion”, sums the thesis up and contains final remarks.

CONTENTS

LIST OF TABLES, FIGURES, AND BOXES

Tables

Figures

"How-to" boxes

CHAPTER 1. INTRODUCTION

Modeling spoken word recognition

Word recognition (or lexical access, as it is also called interchangeably) is one of the most fundamental topics that are being studied in adult cognitive psycholinguistics. For the auditory modality, it concerns the question of how the human mind is capable of perceiving individual spoken words and retrieving their form in real time, be the words listened to in isolation (i.e. separately), or be they heard connected into continuously running speech. Experimental research on spoken word recognition reaches back over a whole century already (see e.g. Bagley, 1900, for a very early study) and has since become a large and intense field with its own behavioral studies, laboratory techniques, and paradigms (see the guide by Grosjean & Frauenfelder, 1997). It has been found that the precise manner of how spoken words are being successfully looked up in the mental lexicon—the knowledge repository for words, and their subunits and properties, that hypothetically resides in our brain—is quite a complex cognitive process, but one that operates with an efficiency unmatched, and seemingly without any effort (recent overviews on spoken word recognition studies include Dahan & Magnuson, 2006; Jusczyk & Luce, 2002; Libben & Jarema, 2002; McQueen & Cutler, 2001; Warren, 2013, among others).

In order to explain and consolidate the plentiful empirical results that have accumulated over time, a range of models of spoken word recognition have been proposed. At the beginning, theorizing and modeling was verbal. That is, the pioneering early models described and explained the process in the form of a text, even if that was combined, more often than not, with some kind of schematic illustration (e.g. a box-and-arrows diagram, presenting a number of processing steps, or showing the flow of information through the system). Representative, and most prominent, examples of verbal models include Morton's (1969, 1970) Logogen model, Forster's (1976) Search model, Marslen-Wilson and colleagues' Cohort model (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987), Klatt's (1979) lexical access from spectra (LAFS) model, Cutler and Norris's (1979) Race model, as well as Grosjean and Gee's (1987) prosodic structure model. Besides them, there also exist mathematically oriented theories, such as the fuzzy logical model of perception (FLMP; Oden & Massaro, 1978; Massaro, 1989) or the neighborhood activation model (NAM; Luce & Pisoni, 1998; Luce, Pisoni, & Goldinger, 1990).

The most recent generation of models, which is making a strong impact on current psycholinguistics, has taken a radically different approach, however. Computational models have been built, which pursue the goal of dynamically simulating on a computer, as closely and psychologically correctly as possible, how we humans process language and speech. These models of the mind (Boden, 1988) are part of an emerging new discipline, which has been called computational psycholinguistics—a highly interdisciplinary research field at the crossroads of linguistics, psychology, and computer science (as shown in Figure 1 in the form of a Venn diagram). Computational psycholinguistics is now represented by a large number of dedicated textbooks and general publications that have appeared in the last two decades (Altmann, 1990, 1997; Christiansen & Chater, 2001; Crocker, 1996; Crocker, Pickering, & Clifton, 1999; Dijkstra & de Smedt, 1996a; Ellis & Humphreys, 1999; Grainger & Jacobs, 1998a; O'Reilly & Munakata, 2000; Plunkett & Elman, 1997; Reilly & Sharkey, 1992, etc.; see also Frauenfelder, 1996; Gaskell, 2007; Norris, 2005, for further introduction to the modeling of spoken word recognition in particular).

Computational models have several methodological advantages. While verbal models usually leave by their nature considerable room for being vague, incomplete, or sometimes ambiguous and inconsistent, the computational approach forces the model-building scientist to great accuracy, completeness, and internal coherence (Schade & Berg, 1992). Quite simply, a simulation model cannot run on the computer and cannot produce any results, unless all

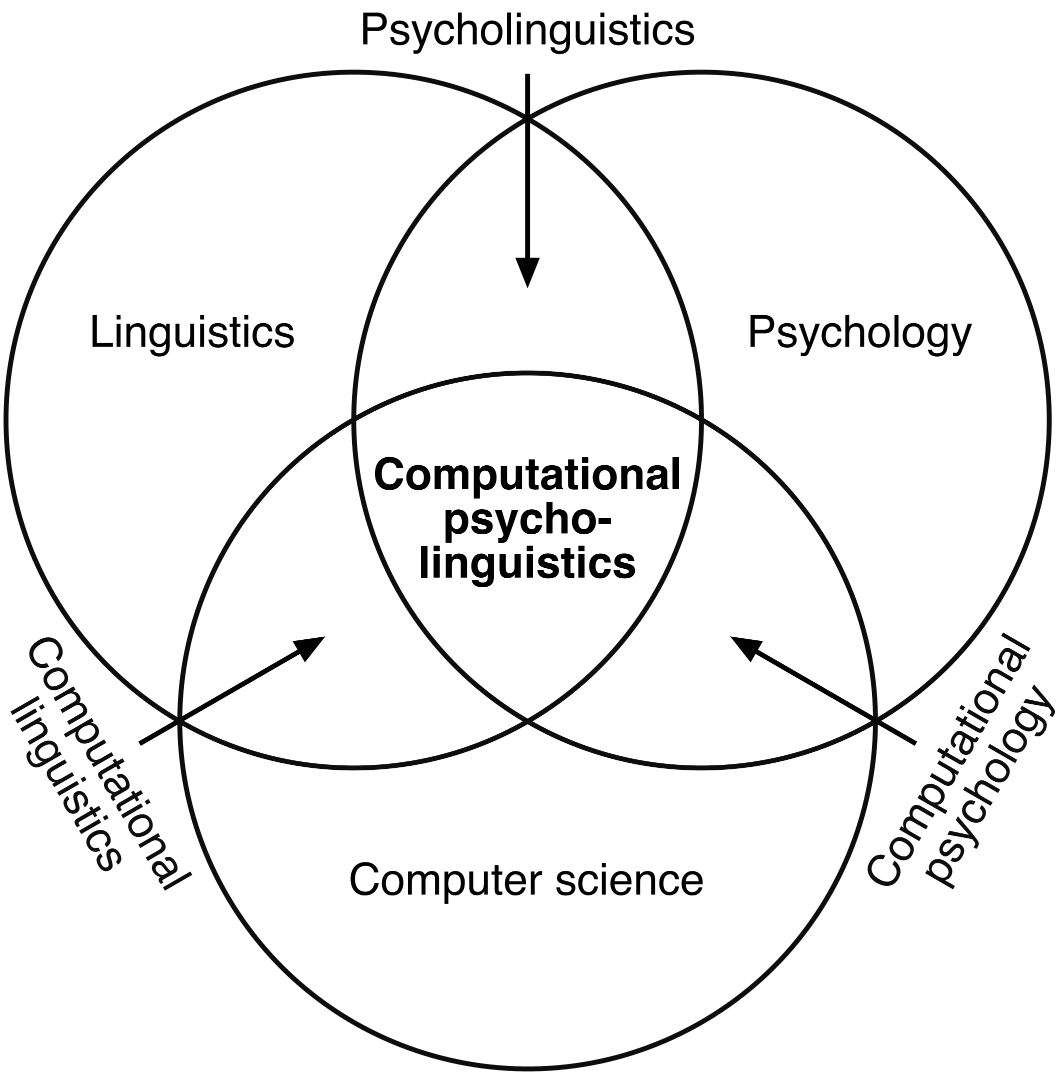


Figure 1. Computational psycholinguistics as the center of a Venn diagram intersection of the disciplines of linguistics, psychology, and computer science.

its internal mechanisms and various knowledge representation bases are clearly and firmly stated (and are actually programmed by means of some computer algorithms and data structures; see Wirth, 1986). In addition, if appropriate graphics methods are implemented, computational models may offer the possibility to visualize the cognitive process under scrutiny. Charts showing the model's state at some moment, or its progress over time, can be displayed on screen, which are more directly informative, and easier to comprehend, than long lists of raw numbers that have to be analyzed outside the model. On condition that the simulation program runs fast enough, these displays may be assembled on the fly, that is, dynamically, while the model is running. As an extra benefit, some (but of course not all) computational models can be manipulated freely and conveniently by user interaction, and can thus be demonstrated live. These models in particular are very rewarding for an educational use (e.g. in a course in psycholinguistics or language sciences).

On a more abstract level, computational models allow for the important, mutual feedback between empirical research and theoretical formalization. This is visualized schematically in Figure 2 (to be read in clockwise direction, starting from the upper left of the figure). The computational psycholinguist devises theories that account for a certain number of experimental results, she or he builds them as computational models, then runs these models on the computer so as to obtain simulation results, and finally relates the simulation results to the original experimental data. Computational models may even make interesting unexpected predictions, which can of course be tested experimentally too. Over time, a continuing research cycle that consists of evaluating and corroborating (or maybe refuting) computational models by more empirical findings and, vice versa, of explaining experimental results by new computational models, will yield to an always more accurate understanding of the psycholinguistic process under study and the various factors it involves (Dijkstra & de Smedt, 1996b; Roelofs, 2005).

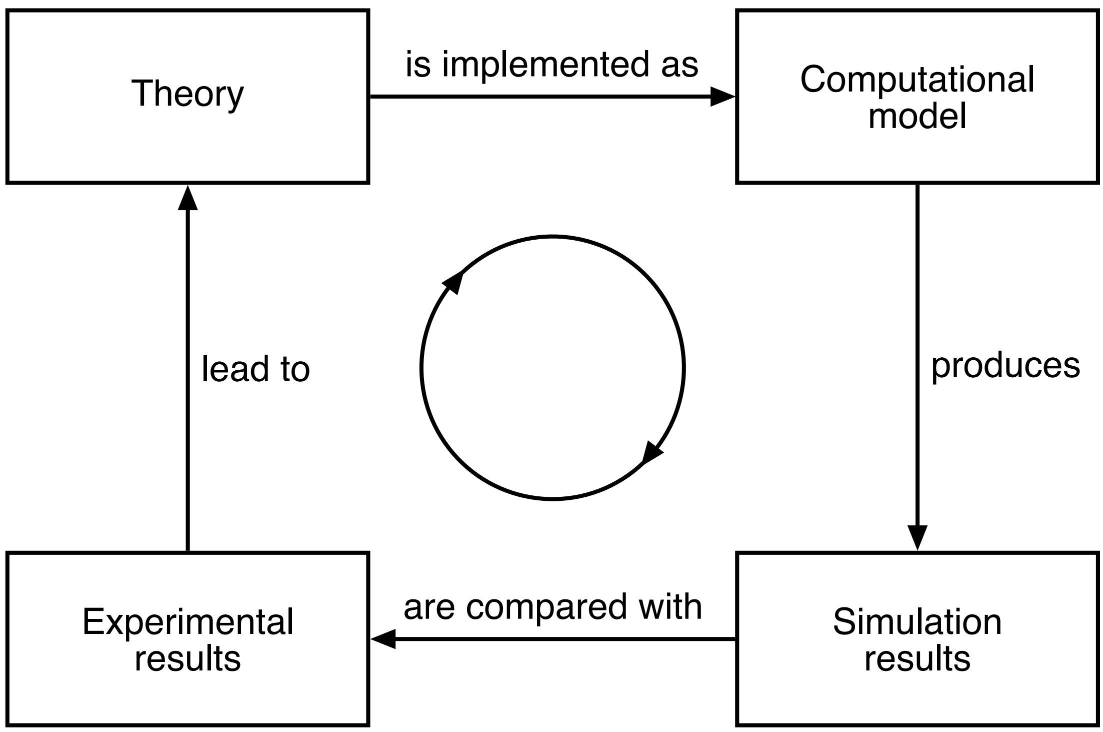


Figure 2. The research cycle in computational psycholinguistics.

Previous models

One of the very first computational models proposed in the formal approach that we will follow in this thesis is the interactive activation model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) of visual word recognition. This model introduced many of the key concepts of localist connectionism, but it dealt with written words and letters. An extension of it to bilingual processing, was proposed by way of the BIA model (Dijkstra & van Heuven, 1998; van Heuven, Dijkstra, & Grainger, 1998). It includes language nodes (one for each of the two languages, English and Dutch), which sum up the activity of the words in the respective lexicon and which repress, by means of inhibitory links, the words of the other lexicon. (A later version, called BIA+ and introduced by Dijkstra & van Heuven, 2002, has so far remained a purely verbal model.).

As concerns the modality of speech, the TRACE model (McClelland & Elman, 1986) has been extraordinarily influential. A localist connectionist model as well, it consists of three levels of units, representing features, phonemes, and words, respectively, and is characterized by activation, competition, as well as interaction. Patterns of activations enter the feature units, and are then propagated, via a number of permanent connections (activatory ones between all levels, and inhibitory ones within the phoneme and word levels), to appropriate phoneme and word units, and also back down again from words to phonemes. It has been shown that, as a result of this ingenious architecture, TRACE is capable of simulating a large number of experimental findings, pertaining both to the identification of phonemes and to the recognition of words (notably, sequences of several words). Besides the authors' own evaluations, TRACE was assessed in a number of follow-up simulation studies carried out by other researchers, most of the time with very favorable outcome (e.g. Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Magnuson, & Tanenhaus, 2001; Frauenfelder & Peeters, 1990, 1998; but cf. also Marslen-Wilson & Warren, 1994). Some of TRACE's characteristics have also met with reservation or been questioned. For one, TRACE contains no more than 14 phonemes, which limits the type of words that can be accounted for, and it has only between 212 and 1,024 words (the exact number depends on the simulation). The small lexicon size is closely tied to the much discussed (especially by Norris, 1994) and now notorious fact that all the word units in TRACE are reduplicated over time (i.e. there is a separate, independent unit for each potential position of a word). This, apart from being rather implausible psychologically, slows down considerably any simulation one might wish to do with larger lexicons (Frauenfelder, 1996; Strauss, Harris, & Magnuson, 2007).

Shortlist (Norris, 1994) is another important computational model for spoken word recognition (two revised versions have been proposed in Norris, McQueen, & Cutler, 1995; Norris, McQueen, Cutler, & Butterfield, 1997). It has a word level, containing inhibitory links similar to those in TRACE, but no feature or phoneme levels. In Shortlist, words become activated directly by means of a search through an electronic dictionary file, combined with the use of an arithmetic scoring procedure, which adds up +1 for each phoneme of a word that matches the string of phonemes supplied to the model and −3 for each phoneme that mismatches it (this can be regarded as a rather rough form of phoneme to word activation and phoneme to word inhibition, respectively). Only the candidate words that match the input string best, that is, a shortlist limited to 30 words per input segment maximum, are involved in determining, by means of competition, the finally winning candidates. All the other words—obviously, the lion's share of the dictionary—are not allowed to have any influence on this competition; this is in contrast to TRACE which keeps all its words involved, at least potentially, from the very beginning to the very end of the word recognition process. As a consequence, even though Shortlist is said to be able technically to draw on dictionaries of 6,000 to over 26,000 words (which signifies a great advantage over TRACE indeed), the model has actually, at any given moment in time, very few words active: 30 per segment, a rigid number that we consider too small, too constraining, and entirely arbitrary. Also, whether a word is a member of the shortlist or not, and if indeed, which bottom-up score that word receives, is being updated only when a new input phoneme arrives (rather than repeatedly and progressively as in TRACE), which is a further unnecessary restriction.

PARSYN (Luce, Goldinger, Auer, & Vitevitch, 2000; see also Auer & Luce, 2005) is a localist connectionist model composed of three levels like TRACE, and pays particular attention to the role of phonotactics in spoken word recognition. Instead of features and phonemes, PARSYN contains two levels of position-specific allophones. One level serves to receive and pass on the input to the model, the other level is used to represent phonotactic probabilities by means of higher resting levels (for allophones that are more often found in that particular temporal position) and within-level activatory connections (both forward and backward within pairs of temporally adjacent allophones that frequently occur together). There are 38 allophones in an initial temporal position and duplicate sets of 50 allophones in three subsequent positions (the first position has fewer units since some allophones cannot appear there), making a total of four positions altogether. Evidently, phonotactic properties are encoded very explicitly in PARSYN, while in other models (e.g. TRACE), they are assumed to emerge implicitly from the lexicon, at least to some extent. We should also note that PARSYN is able to process single short words only (four segments at most). Contrary to TRACE and Shortlist, PARSYN cannot recognize longer words nor sequences of words; therefore the question whether to reduplicate word units, or to use a shortlist, does not present itself.

Next, we mention two computational models that have explored views of spoken word recognition that differ somewhat from the localist one described so far. The Distributed Cohort model (DCM; Gaskell & Marslen-Wilson, 1997) is implemented as a simple recurrent network (Elman, 1990) that was trained to map from a level of features, via hidden and context units (which recall previous sequential states of the network and thereby allow to neatly sidestep the issue of unit reduplication over time), to a level of distributed patterns representing words. There are no phonemes, allophones, or other sublexical units. When several candidate words are compatible with the features that have been input so far, the DCM does not activate multiple independent word nodes in parallel (as it happens in a localist model), but rather it produces a blend on the word level, that is, one single aggregate pattern midway between all the compatible words. (There is an upper limit on the number of words that can be usefully blended together before it becomes too difficult to make a distinction between active and inactive words; see Gaskell & Marslen-Wilson, 1999.) Yet another perspective on spoken word recognition is offered by ARTWORD (Grossberg & Myers, 2000), a model based on the ART framework of adaptive resonance theory (Grossberg, 1976) and described by a set of differential equations; simulations are performed by numerical solution using the MATLAB software. The ARTWORD model defines speech events in real time (i.e. in ms) and can automatically adjust itself to a speech rate according to surrounding context (see the gain control mechanism described by Grossberg, Boardman, & Cohen, 1997). ARTWORD postulates a working memory level representing solitary phonemic items, as well as a chunk level that stores lists of variable length that are activated from grouping these items together to words. Longer chunks mask shorter ones, which is comparable to lateral word inhibition in TRACE, Shortlist, and PARSYN. When a chunk wins, it reinforces the items that activated it—this bears a resemblance to TRACE's word to phoneme activation; doing so, it creates the ART framework's typical resonances. (The simulations reported by Grossberg & Myers, 2000, used no more than four words and eight phonemes. It is unclear if ARTWORD can be scaled up to a large lexicon.)

In the last few years, the Bayesian theoretical framework has become quite popular in cognitive modeling, which has led Norris and McQueen (2008) to reformulate Shortlist in Bayesian instead of connectionist terms. Shortlist B, as the model is now called, basically calculates a chain of probability estimates: probabilities of phonemes (as derived from a confusion matrix of two-phoneme sequences), products of probabilities of the phonemes in a word (for the words that begin at a particular segment), prior probabilities of words (as derived from a word frequency database), probabilities of paths (i.e. ways of segmenting the sequence of phonemes into sequences of words), and posterior probabilities of words given the evidence. In Shortlist B, as in the original model, the number of word candidates being considered at any moment in time is explicitly limited (50 per segment), as are the number of paths (500 in total). Even though Shortlist B was promoted by its authors as a radical departure from connectionism, others have explained that the Bayesian and connectionist frameworks are not really incompatible (McClelland, 2013; McClelland, Mirman, Bolger, & Khaitan, 2014).

We human beings are experts in hearing an acoustic wave carrying speech and processing it directly. A full model of the psycholinguistic process would therefore not only simulate the recognition of spoken words but it would also account for all that is involved in perceiving and analyzing real speech, and categorizing it into the preliminary sublexical representations (be they features, allophones, phonemes, clusters, etc.) that are at the foundation of spoken word recognition (Liberman, 1996; J. Miller, 1990; Moore, Tyler, & Marslen-Wilson, 2009; Pisoni & Remez, 2005; Raphael, Borden, & Harris, 2011). All six models we have described start instead from an explicit, abstract encoding of speech (McQueen, 2007): mock-speech feature patterns in the case of TRACE and the DCM; vectors describing subjective similarities among allophones in PARSYN; sequences of phonemes (or their likelihoods) in Shortlist, Shortlist B, and ARTWORD. Our models will follow a similar approach. But a totally different approach, represented by models like RAW (van Kuijk, Wittenburg, & Dijkstra, 1996) and SpeM (Scharenborg, Norris, ten Bosch, & McQueen, 2005), is to integrate some techniques developed for automatic speech recognition (see e.g. Calliope, 1989; Huang, Ariki, & Jack, 1990; Kohonen, 1988; Lee, 1989; Lippmann, 1989, 1997; Pieraccini, 2012; Rabiner & Juang, 1993; Waibel & Lee, 1990, etc.) as front-ends to models of human spoken word recognition. For now, it remains to be seen if such activities will be fruitful in the long run to understand human spoken word recognition: by their nature, these engineering methods tend not to take into account much of what is known about the psycholinguistic process (see the critique by Massaro, 1996).

CHAPTER 2. TWO NEW MODELS: FN5 AND BIMOLA

Although the existing models have undeniable value and importance, we see good reasons to propose two new models of spoken word recognition. One of our models goes under the name of FN5 and simulates French monolingual spoken word recognition. It recognizes single (i.e. isolated) words as well as sequences of connected words. The other model is called BIMOLA, which is an acronym for bilingual model of lexical access, and accounts for English–French bilingual spoken word recognition. While BIMOLA deals with single words only, these words can come from one language (the bilingual's base language) or from the other (the guest language).[1]

Reasons to propose two new models

1. To simulate spoken word recognition in French. All the spoken word recognition models we have mentioned in the previous chapter are concerned exclusively with English (Shortlist B with Dutch). While some operations of the

[1] We described an earlier version of FN5 in Léwy, Grosjean, Grosjean, Racine, and Yersin (2005) and an earlier version of BIMOLA in Léwy and Grosjean (2008).

human language faculty can safely be considered universal to all the world's tongues, many if not most psycholinguistic processes must take into account a great deal of language-specific aspects, such as the content of the lexicon, the various properties of its words, the repertoire of sublexical units that exist, etc., which of course differ considerably from one language to another (Cutler, 1997, 2012). As a consequence, there is a pressing need to establish computational models for languages other than English. Contributing to this cause, our model FN5 is the first to simulate spoken word recognition in French.

2. To model phonological phenomena within words and at boundaries of words. Although several existing computational models are capable of accurately identifying spoken words when they get the sequence of sublexical segments (e.g. phonemes in TRACE and Shortlist, or allophones in PARSYN) that constitute these words in their canonical form, the models generally do not simulate phonological phenomena that can take place within a word or at boundaries between words, and which change the words' pronunciation. Our model FN5 deals with a number of phonological phenomena that are central to French (deletion of schwa, linking with and without liaison, word contractions including elision), using a representation method that is also suitable for other languages.

3. To introduce a new approach to recognizing sequences of words. Two of the most influential localist connectionist models, TRACE and Shortlist, present certain architectural drawbacks in regard to how they allow for the recognition of multiple spoken words presented sequentially. TRACE freely lets all its words participate in the process at all times but, to do so, needs to reduplicate word units over all potential positions, including many unlikely ones. Shortlist, by contrast, does without multiple copies of words but it has to restrict the number of words that are permitted to take part in the process at any given moment to just a few. Since neither of the two solutions is ideal, our model FN5 introduces a new approach (and compromise), in which as many words can be active simultaneously as in TRACE, but which functions without reduplication like Shortlist.

4. To account for differences between dialects. Former models cover the standard language only and do not take into account any regional and dialectal characteristics (e.g. the idiosyncrasies that distinguish American from British English, not to mention those of the various varieties within either group). We touch in the model FN5 upon a few salient differences between standard (i.e. Parisian) French and Swiss French, in terms of which are the phonemes that exist and how some of the words are pronounced differently. We can make this model operate in either one of these two varieties of French, and we can draw comparisons.

5. To simulate spoken word recognition in bilinguals, particularly the recognition of guest words. Other than the written word recognition model BIA, all the models mentioned are monolingual. As for spoken word recognition, an extensive body of experimental results on bilinguals has accrued (Grosjean, 1988, 1997, 1998, 2008; Li, 1996; Soares & Grosjean, 1984) and calls for a computational model. This is precisely the role of BIMOLA, our second model. BIMOLA accounts for the bilinguals' ability to process auditory words from either of two languages (currently English and French), to go in and out of various language modes (bilingual vs. monolingual), and to recognize guest words, that is, code-switches and borrowings from one into the other language.

6. To advocate simultaneous activation of two languages in bilinguals. Through various activatory mechanisms running in both languages entirely independently yet in parallel, BIMOLA can do away with language nodes (as proposed by BIA). There is no empirical evidence that such nodes exist, nor do we know how a new node is created when a language is learned. In BIMOLA, language activation is distributed over all the nodes (phonemes and words) of a language and hence a specific node dedicated to this purpose is not necessary. Inhibition between languages (as postulated by BIA) is absent from BIMOLA, since it typically makes one of the two languages end up being deactivated by the other language. By contrast, in BIMOLA, both languages are active, albeit one more than the other.

7. To use extensive sets of words and phonemes. Apart from revised versions of Shortlist (Norris et al., 1995, 1997) and from Shortlist B, previous computational models draw on heavily reduced bases of linguistic information. That is, these models typically use some small-sized lexicon (e.g. 212 to 1,024 words in TRACE, and even just four word chunks in ARTWORD); also, they often operate with an incomplete set of sublexical units and positions (e.g. 14 phonemes in TRACE, and four positions of allophones in PARSYN). In contrast, each of our own two models contains a substantial lexicon: there are 17,668 words, some of them with variants, hence giving rise to a total of 20,523 pronunciations, in FN5; and there are 8,696 words, 4,348 for each language, in BIMOLA. Furthermore, both models possess a full inventory of phonemes (standard French or Swiss French in the case of FN5, and standard French and English for BIMOLA), employed in as many phoneme positions as appropriate.

8. To represent similarities and differences of phonemes within and between languages. Computational models of spoken word recognition use very different encoding systems to numerically represent the speech input (e.g. feature patterns in the case of TRACE and the DCM, vectors of subjective similarities among allophones in PARSYN, etc.), but none of these encoding systems describes more than a single language. We prepared a bilingual/ bidialectal phonetic feature matrix, which is used in the two models. It covers all the phonemes of English and French (including Swiss French) and allows to quantify, with the help of a metric, the similarities and differences of phonemes both within each language and between the two languages.

9. To study various speech rates. In most prior models (all mentioned with the exception of ARTWORD), the process of temporally unfolding a spoken word from beginning to end (i.e. from the word's first phoneme to its last phoneme) takes place at one permanent speech rate. While we will put a normal speech rate into use for the majority of the simulations, we will also show how the speech rate (and hence the word's temporal structure) can be varied in our models in principle, and what effect a faster speech rate has in practical terms.

10. To run simulations easily and comfortably. Not all implemented computational models include a user interface along with their core simulation software. Those that do are generally easier and more comfortable to use and manipulate (notably some implementations of TRACE; see Warren, 1993; Strauss et al., 2007). Those without user interface have typically been run by their authors only and have rarely been released to the scientific community. Our models offer a modern and intuitive graphical user interface, switchable between English and French; they run on Apple Macintosh computers (OS X) and are simple to install and use. So, parties interested in doing simulations on their own (e.g. in context of new experimental research), or using one or both models for teaching, could do that with no trouble at all.

Architecture of FN5 and BIMOLA

Both the FN5 model and the BIMOLA model are couched in the widely known formalism of localist connectionist networks, which was originally introduced by the interactive activation model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982).

In a localist connectionist model, every single of the processing units (also called nodes), which are usually organized in some hierarchy of levels, stands for one particular abstraction about the model's input. Each such unit can be activated at the same time as and separately from all the other units—much like a huge cluster of light bulbs that can be turned on simultaneously and be brightened up or dimmed down individually.[2] For just this reason, a localist connectionist model is able to represent, in parallel, a very large number of independent abstractions (possibly totally different ones), all at their appropriate level of activation, and can draw a remarkable power from the units' complex yet systematic interaction and evolution over time. Connections between units, responsible for this behavior, are not trained automatically but rather judiciously set by hand (i.e. by the person who builds the model), and are usually weighted in a uniform manner within each class of connections (see Grainger & Jacobs, 1998b; McClelland & Rumelhart, 1988).

If localist connectionism is today, on the whole, a little less common than the usual type of connectionism that relies on distributed representations and back-propagation learning (Rumelhart, Hinton, & Williams, 1986), it has all the same continued steadily to attract quite a few researchers in the psychological modeling community, both inside and outside the area of language. Apart from the localist connectionist models already mentioned so far, we call attention to those presented in the volume edited by Grainger and Jacobs (1998a) and to several models that have been proposed individually (see e.g. Berg & Schade, 1992, 2000; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; G. Dell, 1986, 1990; Feldman & Ballard, 1982; Gilbert & Shallice, 2002; Goldstone, 1994; Grainger & Jacobs, 1996; Harley, 1993; Jacobs & Grainger, 1992; McClelland, 1986, 1991; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Norris, McQueen, & Cutler, 2000; Page, 2000; Shastri & Ajjanagadde, 1993; Stemberger, 1985; Waltz & Pollack, 1985, among others).

The general architecture of FN5 is shown in Figure 3. There is an organization into three levels of localist connectionist units: features, phonemes, and words (for specifics of the linguistic knowledge that is present on each of these levels, refer to the next chapter). We find bottom-up connections running from features via phonemes to words, then top-down connections from words back to phonemes (they are optional as one can see in the figure), as well as lateral connections within both the phoneme and word levels. These links may be activatory or inhibitory, which is shown by their ending in arrows or circles, respectively, as is the usual custom. Manifestly, this architecture is informed and inspired partly by previous models of spoken word recognition, and this in particular by TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994),

---

[2] Jeff Elman (personal communication, 1996) suggested to us this characterization.

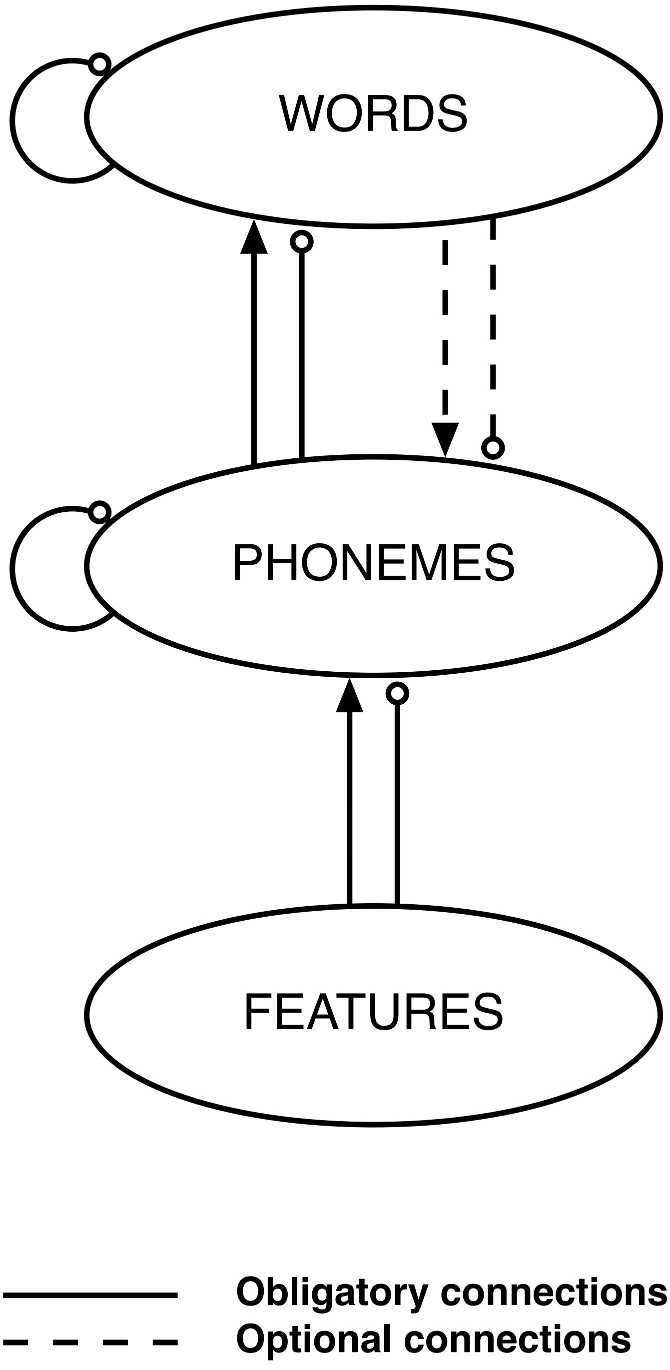


Figure 3. General architecture of the FN5 model.

both localist connectionist models too. With its three linguistic levels (the same ones as in TRACE), FN5 dissociates itself from direct mapping models that operate without any intermediate sublexical representation, such as the LAFS model (Klatt, 1979) and the DCM (Gaskell & Marslen-Wilson, 1997). It equally sets itself apart from models that have additional representation levels as, for example, a level of allophone co-occurrences like PARSYN (Luce et al., 2000).

Concerning the flow of information between the three levels, we draw attention to the presence of bottom-up inhibition (as in Shortlist but not in TRACE) and to the option of top-down feedback (present in TRACE but not in Shortlist).

Bottom-up inhibition is contained in several spoken word recognition models. It can be categorical, that is, as soon as a mismatch (even only a small one) is detected between the arriving input and a potential candidate word, that candidate is eliminated at once. Such it was the case in the original version of the Cohort model (Marslen-Wilson & Welsh, 1978). But the inhibition can be also be gradual, in the way that the word candidate's level of activation is merely reduced by some degree (not totally), as it was proposed in the revised Cohort model (Marslen-Wilson, 1987) and as it is also implemented in Shortlist's scoring procedure (which we have already mentioned). In divergence to these models, TRACE did not include any bottom-up inhibition at all; McClelland and Elman (1986, p. 55f) argued that the same effect was lexically mediated by lateral inhibition between competing word candidates. So, for example, when an input pattern that stands for "pleasant", is entered in TRACE, the word "present" is inhibited, not by bottom-up inhibition (from phoneme /l/ to "present") but rather by lateral competitive inhibition (from the more highly active word unit of "pleasant" to the less highly active word unit of "present"). For TRACE, when no better word candidate exists in the lexicon, a mismatch occurring in a non-initial position has no consequence. While results from some earlier experimental studies (e.g. Marslen-Wilson & Warren, 1994; Zwitserlood, 1989) did not allow to clearly discriminate between direct (i.e. bottom-up, between-level) vs. mediated (lateral, same-level) inhibition effects, subsequent findings such as those by Frauenfelder, Scholten, and Content (2001) made progress in that direction, and confirmed the importance of an independent component of bottom-up inhibition. We therefore have implemented it (but in a rather new form, as we will explain when we come back to it below for details). In our opinion, lateral inhibition and bottom-up inhibition take on complimentary roles in spoken word recognition: Lateral inhibition is responsible for causing the best matching word to win slowly but surely over other strong candidates (cf. McClelland & Rumelhart's (1981) rich-get-richer effect), whereas bottom-up inhibition serves to rule out obviously mismatching words quickly and early on in the matching process (even so if this means that there is no other, better candidate and therefore possibly no winner at the end).

As for the long-running debate on the presence or absence of top-down feedback from words to phonemes remains, it remains, for the time being, very lively (for a review, see McClelland, Mirman, & Holt, 2006). There are the ones that are in support of an interactive approach to speech processing as it was instantiated by TRACE (e.g. Elman & McClelland, 1988; Magnuson, McMurray, Tanenhaus, & Aslin, 2003; Pitt & Samuel, 1995; Samuel & Pitt, 2003), and there are the others that strictly adhere to an autonomous approach (e.g. Massaro, 1989; and particularly Norris et al., 2000, in their Merge model on phoneme decision). Both camps have continued to advance empirical evidence and theoretical arguments, also from the angle of how speech perception may be influenced by training, adaptation, and attention (Norris, McQueen, & Cutler, 2003; Mirman, McClelland, Holt, & Magnuson, 2008). At the sight of this issue still being under considerable dispute and requiring further experiments until it can be resolved, we have decided to adopt a neutral stance: We do have included connections from words to phonemes but we have made them optional (as shown by the dashed lines in the figure). Normally, the two mechanisms, top-down activation and inhibition, are turned off in FN5; they have remained so for all the simulations we will report (these simulations are concerned with the identification of words rather than phonemes, anyway). However, we offer the possibility to switch top-down feedback on, should one wish to do so.

Figure 4 presents the general architecture of BIMOLA. As can be seen, BIMOLA basically consists of the same three linguistic levels as FN5, that is, features, phonemes, and words. The feature level units are identical to the features of FN5 and meant to be extracted from the acoustic wave; they are shared by the two languages. Phoneme and word units, by contrast, are organized according to the subset hypothesis proposed by Paradis (1989), that is, by itself (each language is represented by a subset of units) but also as one large system (both subsets are enclosed in a larger set).[3] Connections (bottom-up, top-down, and lateral) are essentially the same in BIMOLA as in FN5, but they exist in BIMOLA, as shown in the figure, separately within each language (A and B) and thus form two language networks: one is leading from features via language A phonemes to language A words, the other language network goes from the features via the language B phonemes to the language B words. The two networks operate in parallel (i.e. at the same time) and run entirely independently from each other (there is, in particular, no inhibition from one language to the other). At both the phoneme and word levels, units can have near or distant neighbors, which are visualized in the figure by how dark and

[3] The language pair that we have implemented is English and French, but BIMOLA is meant to be a general bilingual model and independent of language pair; hence the figure uses the labels, language A and language B.

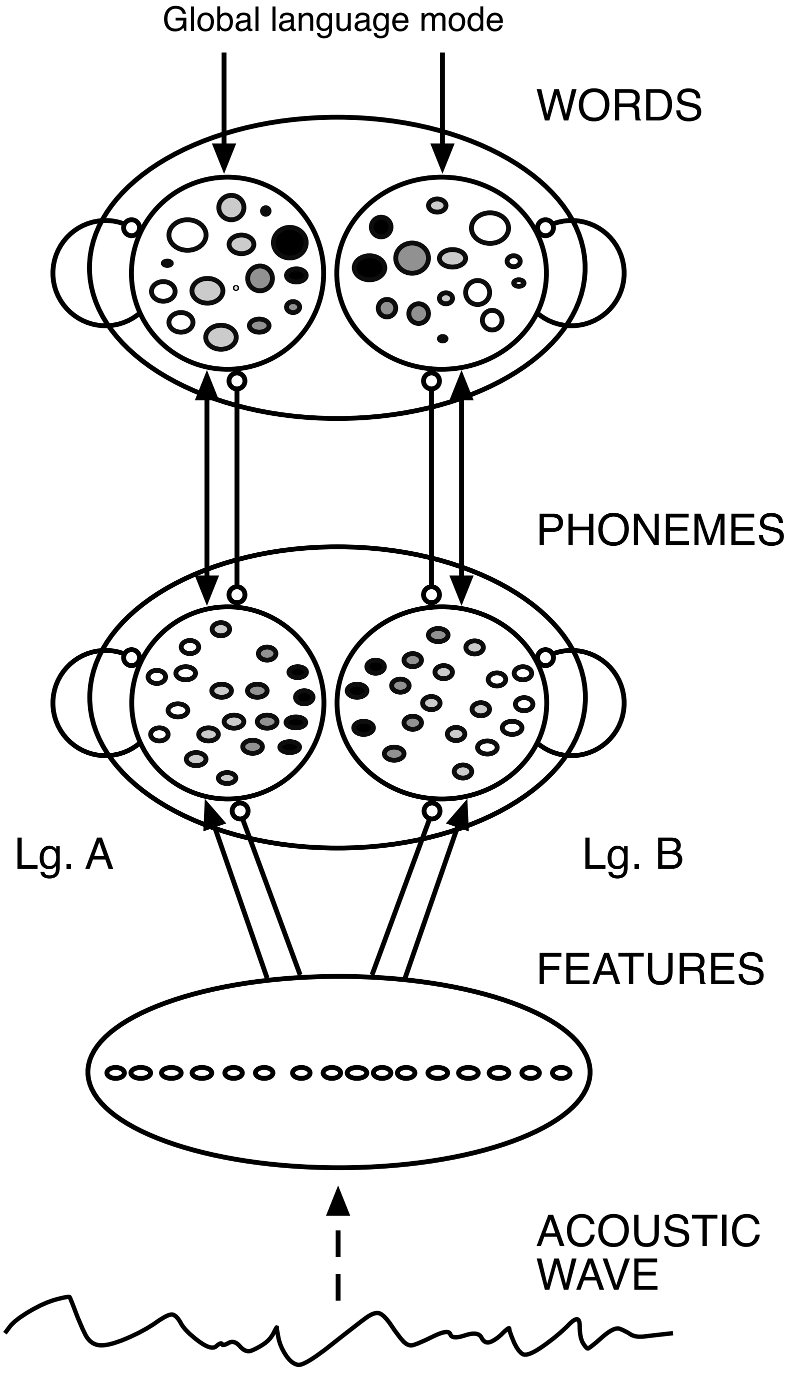


Figure 4. General architecture of the BIMOLA model.

how close they are spatially. At the word level, frequency is accounted for; this is depicted by the units' size.

As shown at the very top of the figure, words and phonemes are pre-activated based on information on the listener's global language mode. At any moment of time, one of the two languages of a bilingual is usually the main language of listening; this is called the base language. Which one is it presently, language A or language B? At the same time, the other language, called the guest language, will be activated to various degrees as well. It can be highly activated in case of a bilingual language mode, or it can be only a little activated in the event of a monolingual language mode. To which degree is it activated currently? BIMOLA can be operated with either French or English as the base language, and the guest language may be activated anything from 100% down to 0%. Of course, in the last case, BIMOLA runs with only one language (i.e. one set of phonemes and one set of words) and becomes a purely monolingual model, quite similar to our other model, FN5. For the most part, we will examine monolingual simulations with English words (English as a base language, French inactive) as well as bilingual simulations (French as a base language, English as a guest language) using code-switches and borrowings from English into French.

Remaining chapters of this thesis

The remaining chapters of this thesis give the particulars of both our models, FN5 and BIMOLA, and will in turn establish the two models' linguistic knowledge (Chapter 3), discuss their internal mechanisms (Chapters 4 and 5), and present their systematic evaluations (Chapters 6 and 7). As will become apparent, there is a substantial area of theoretical, methodological, and implementational overlap between FN5 and BIMOLA, and all the differences between the two models are purely motivated by their having a different focus: the French language, isolated and connected word recognition, pronunciation variants, etc., for FN5; bilingualism, accessing words in English and in French, processing code-switches and borrowings, in the case of BIMOLA. At this point, it is therefore useful to clarify which parts of this dissertation are common to both models and which are specific to either one or the other model. Table 1 provides exactly this information.

Table 1.
Elements of this thesis common to both models and specific to one or the other

Common to both models	
Formalism and architecture: - Localist connectionist network	 see this chapter
Sublexical information: - Phonetic feature matrix - Metric space of phonemes	 see Chapter 3, "Features" see Chapter 3, "Phonemes"
General mechanisms: - Phoneme activation and inhibition - Normal vs. fast speech rate - Word activation, inhibition, and isolation	see Chapter 4
Specific to FN5	
Lexicon: - French nouns, determiners, and adjectives (used in either standard or Swiss version)	see Chapter 3, "Words for FN5"
Specific mechanisms: - Position and variant selection - Attenuation for ending variants - Influence of the context	see Chapter 5, "Sequential processing in FN5"
Evaluations: - Simulations on isolated words - Simulations on connected words	see Chapter 6
Specific to BIMOLA	
Lexicon: - English and French verbs (both languages used or one language alone)	see Chapter 3, "Words for BIMOLA"
Specific mechanisms: - Global language mode - Recognizing guest words, with or without guest language pronunciation	see Chapter 5, "Language activation in BIMOLA"
Evaluations: - Monolingual simulations - Bilingual simulations (base language words, code-switches, and borrowings)	see Chapter 7

CHAPTER 3. LINGUISTIC KNOWLEDGE

FN5 and BIMOLA rely on rich, detailed information at each of their three levels of linguistic description: the features, phonemes, and words. Taken together, these representations constitute the models' knowledge about concrete human languages. FN5 covers two versions (i.e. geographical varieties or dialects) of French: standard French and Swiss French. Depending on the lexicon loaded into the model at run-time, one or the other of two alternative sets of phonemes is selected (we will indicate which phonemes in a moment). BIMOLA deals with standard French and English (here too we will provide the details further down); it can be run as a monolingual model (when only the French part or only the English part of the model is used) and it can operate as a bilingual model (if the French and English parts are used both together). We begin by presenting the features, which are the same in FN5 and BIMOLA, we then continue with the phonemes and present a distance measure and metric space of phonemes, and we finish with separate sections for each model on the words and lexicons.[4]

---

[4] In principle, both models could be adapted to or replaced with linguistic knowledge describing other languages, and the models would essentially still function. But it takes a great deal of time-consuming, meticulous work to prepare such knowledge bases, and new ones would have to correspond in content and format to ours.

Features

The two models use the same phonetic feature matrix, which we have originally put together for BIMOLA for English and standard French,[5] and which we have subsequently adapted for FN5 to Swiss French. In Appendix A, it is given in full, over four pages. A feature matrix is basically a large table with phonemes as rows and their feature values as columns. Prior to our proposal, such feature matrices were available for different languages but invariably described a single language only; none dealt with several (i.e. two or more) languages or dialects at the same time. As monolingual feature matrices do not distinguish phonemes across languages, one cannot simply integrate two monolingual matrices into one to obtain a bilingual matrix. Our feature matrix, which is bilingual (English and French) as well as bidialectal (standard and Swiss French), represents the similarities and differences of phonemes both within each language or dialect, and between the languages and dialects. It is primarily based on the traditional +/− framework of generative phonology, described in SPE (Chomsky & Halle, 1968/1991; French table by F. Dell, 1985), but extends it to a set of 18 features, as presented in Table 2. Some of the SPE features have remained untouched and are employed with categorical, binary values (1 or 0), in the original sense of distinctive features (Jakobson, Fant, & Halle, 1952). But others have been reorganized and have been turned into scalar features (see Flemming, 2001; Ladefoged, 1993), which can take on intermediate values, as shown in the table. A very few features became redundant in that process and have been left out. In addition, three new features have been included; they relate to aspiration (in English plosives), movement (as in English diphthongs and long vowels), and relative length (both within consonants and within vowels), and are also scalar. Scalar features were used previously by at least two monolingual spoken word recognition models: the FLMP (Oden & Massaro, 1978; Massaro, 1989) and TRACE (McClelland & Elman, 1986), which had seven features (mostly taken from Jakobson et al., 1952) with a range of eight values. All the changes we made were necessary to discriminate and compare the phonemes between the two languages and dialects appropriately.

[5] For full details, including discussion of the exact feature values in the matrix, see Léwy, N. (1995). *The phonemes of French and English: Similarities and differences*. Unpublished manuscript, Laboratoire de traitement du langage et de la parole, Université de Neuchâtel, Switzerland. As an alternative to SPE, we also considered Lahiri and Marslen-Wilson (1991), and Shillcock, Lindsey, Levy, and Chater (1992).

Table 2.
The 18 features used in FN5 and BIMOLA

Feature	*Abbreviation*	*Value set*
SPE features that remain unchanged		
Sonorant	[SONO]	0/1
Syllabic	[SYLL]	0/1
Consonantal	[CONS]	0/1
Coronal	[COR]	0/1
Nasal	[NAS]	0/1
Lateral	[LAT]	0/1
Continuant	[CONT]	0/1
Delayed release	[DEL]	0/1
Voiced	[VOI]	0/1
SPE features made scalar and reorganized		
Tense	[TEN]	0/0.5/1
Round	[ROU]	0/0.2/0.5/0.7/1
Low	[LOW]	0–1
Back	[BACK]	0–1
High	[HIGH]	0–1
Front	[FRO]	0–1
New features		
Aspirated	[ASP]	0/0.2/0.5/0.7/1
Long	[LONG]	0–1.5[a]
Moving	[MOV]	0–1

[a] The range of values in BIMOLA is 0–1, and 1.5 is only used for the Swiss French long vowels in FN5.

Phonemes

Standard French (*le français standard*), as it is spoken in the Paris region, has a set of 35 phonemes: /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ, m, n, ɲ, ŋ, l, ʁ, j, w, ɥ, i, y, e, ɛ, ø, œ, ə, a, ɔ, o, u, ɛ̃, ɑ̃, ɔ̃/. This inventory is complete and well established (see e.g. the dictionary *Le Petit Robert*, 1992; and the IPA illustration by Fougeron & Smith, 1999). It reflects that the velar nasal consonant /ŋ/ has, for all practical purposes, become a proper part of the French language, namely for loanwords from English, such as “camping” /kɑ̃piŋ/ (Walter, 1983, 1999). It also accounts for the observation that today most speakers of standard (i.e. Parisian) French do not make a distinction anymore between front /a/ and back /ɑ/, nor between spread /ɛ̃/ and rounded /œ̃/, but usually say /a/ and /ɛ̃/. That means, the words “tache” vs. “tâche” and “brin” vs. “brun” are both pronounced /taʃ/ and /bʁɛ̃/, and are homophones now (see Léon, 1996; Tranel, 1987).

In the Romandie, the French-speaking western part of Switzerland, a variety (or dialect) of French is spoken. Swiss French (*le français de Suisse romande*)[6] is at most linguistic levels very similar to standard French, but at the phonological level, it differs from it on at least two aspects (see Knecht, 1985; Grosjean, Carrard, Godio, Grosjean, & Dommergues, 2007). Firstly, speakers of Swiss French do maintain systematically both the /a, ɑ/ and the /ɛ̃, œ̃/ contrasts. Therefore, “tache” (/taʃ/) and “tâche” (/tɑʃ/), and “brin” (/bʁɛ̃/) and “brun” (/bʁœ̃/), are non-homophones. Secondly, there is an opposition between short and long vowels, which is most obvious in the case of minimal pairs distinguished by the sole duration of the final vowel. The adjective “joli, -ie”, for example, ends in a short /i/ when it is in the masculine form (“joli”), but in a long /iː/ in its feminine form (“jolie”); and for “roux” one says /ʁu/, while unrelated “roue” is pronounced /ʁuː/ (see Métral, 1977). By adding the 7 long vowels that can occur, /iː, yː, eː, ɛː, øː, aː, uː/, as well as /ɑ/ and /œ̃/, to the standard French phonemes listed above, we end up with an extended repertoire of 44 phonemes for Swiss French.

For English, we put forward the following list: /p [pʰ], b, t [tʰ] [ɾ], d, k [kʰ], g, f, v, θ, ð, s, z, ʃ, ʒ, ʧ, ʤ, m, n, ŋ, l [ɫ], ɹ, h, j, w, i, ɪ, e, ɝ, ə, æ, ɑ, ʌ, ɒ, ɔ, ʊ, u, eɪ, aɪ, ɔɪ, əʊ, aʊ, ɪə, ɛə, ʊə/. These 49 sounds include some allophones (aspirated plosives [pʰ], [tʰ], [kʰ], and dark [ɫ]) as we want to differentiate English variants close to and English variants distant from the corresponding French sounds. We use the Received Pronunciation of British English (*Collins COBUILD*, 1987; Gimson, 1989; Roach, 2004), but to address American English to some degree (Ladefoged, 1999), we transcribe the accented central vowel as a rhotacized /ɝ/ (e.g. “work” becomes /wɝk/), and we accept the intervocalic flap [ɾ] (e.g. for the words “tidy” or “fiddle”) as an input to BIMOLA, like any British English phoneme (no actual node for flap is created at the model’s phoneme level and no word in its lexicon, which is based on British English as we will see, contains a flap). To

[6] We consider here Swiss French as a whole even though there are actually several regional varieties, with certain minor differences, within the Romandie (the varieties spoken in Neuchâtel, Geneva, Fribourg, Vaud, the Lower Valais, etc.; Métral, 1977). If ever in doubt, we relied on the Neuchâtel one.

ensure that affricates and diphthongs can be compared with other phonemes, we treat them as single phonemes (e.g. /ʧ/, /eɪ/, etc.) rather than as sequences of two phonemes (e.g. /t/+/ʃ/ and /e/+/ɪ/). To indicate that length is not distinctive for the English vowels, in contrast to the Swiss French vowels, we do not use length marks (e.g. /i/ is always long and /ɪ/ is always short; see Gimson, 1989).

Examples of words that contain the standard French, Swiss French, and English phonemes can be found in the feature matrix (in Appendix A); all those examples are from FN5's and BIMOLA's lexicons. For each sound, we also indicate, in that table's last column, a one-character case-sensitive code, which is fairly self-explanatory (e.g. k for /k/, K for [k^h], 1 for the first diphthong /eɪ/, etc.). It is used to store pronunciations of words in the lexicon files (e.g. "cake" → /k^heɪk/ → "K1k") and can serve to type the phonetic symbol with a keystroke. (This coding is practical and enough for our two models and two languages; for a coding of the complete IPA, see e.g. Kirshenbaum, 2001; Wells, 1995.)

<u>*Distance measure.*</u> As each phoneme is described by a row of values in our feature matrix, we can provide a measure of distance ρ between any two phonemes x and y (i.e. two rows in the feature matrix) by calculating the absolute difference of the feature values and summing it over all features:

$$\rho(x, y) = \sum_{f} (\mid x_f - y_f \mid w_f) \qquad (1)$$

Mathematically, this distance measure is a rectilinear "taxicab" 1-norm (and not the Euclidean 2-norm with squares and roots; Bronstein & Semendjajew, 1991; Krause, 1975/1986). Linguistically, this simplest possible form is appropriate (Goldstone, 1994; Chomsky & Halle, 1968/1991; Jakobson et al., 1952). Within the sum of differences of feature values, the two features [LOW] and [BACK] are weighted with $w_f = 3.8$, the other features with $w_f = 1$. This increased weight has been determined in such a way as to balance the distribution of distances within the 26 English and French vowels (M = 3.85, SD = 1.52, N = 325 pairs) with the distribution of distances within the 44 English and French consonants (M = 3.83, SD = 1.55, N = 946 pairs; $t(1269) = -0.176$, *NS*). Distances for pairs composed of a vowel and a consonant are typically much larger (M = 8.02, SD = 1.30, N = 1,144 pairs; not including the glides, diphthongs, and Swiss French additional vowels).

For practicality's sake, distance ρ is normalized, that is, divided by the maximum distance that is observed over all pairs of phonemes; this maximum is 11.60 and happens to be for French /ɔ̃/ and either French /t/ or English [t^h].

Thus, a normalized distance of phonemes is represented by the formula:

$$\text{dist}(x, y) = \frac{\rho(x, y)}{\max\limits_{(x, y)} \rho(x, y)} \qquad (2)$$

Correspondingly, a normalized proximity of phonemes can be expressed as:

$$\text{prox}(x, y) = 1 - \text{dist}(x, y) \qquad (3)$$

Both functions range from 0 to 1, due to the normalization. To itself, a phoneme has a normalized distance of 0 and a normalized proximity of 1. As we will see when we present general internal mechanisms, these two equations determine, respectively, the phoneme inhibition and the phoneme activation, in our models.

Metric space of phonemes. By measuring geometrically the distance between pairs of phonemes, we introduce a metric over the space of phonemes. In this highly multidimensional abstract space, in which each feature dimension spans an axis, the phonemes have close as well as distant neighbors, according to their perceptual distance (cf. Marslen-Wilson, 1993; Marslen-Wilson, Moss, & van Halen, 1996; see also Auer & Luce, 2005). There are neighbors of phonemes both within a language and across languages. In the bilingual phoneme space, French /b/, for example, is found to be very near to French /d/, /g/, and /p/ (with a distance ρ of 1.2, 1.3, and 1.5, respectively) and to English /b/ (distance of 1.0), but farther away from English /d/, /p/, /t/, and French /n/ (distances of 2.3, 2.6, 3.9, and 4.2), and even more distant to the vowels, French /i/ and English /i/ (distances of 8.0 and 8.5), among others. Such relations exist for all pairs of phonemes. Figure 5 informally illustrates the French /b/ distances, in the upper diagram, and shows distances calculated between English /i/ and various phonemes in the two languages, in the lower diagram. Close neighbors are placed together and are drawn with dark circles; distant neighbors are more spread out and are drawn with clearer circles (in the same way as in BIMOLA's general figure; see p. 17).

To get an even better idea of the distance relations between phonemes, and to examine the metric space of phonemes as a whole, it is instructive to do a hierarchical clustering analysis (cf. Shepard, 1980). For 93 sounds altogether (49 English + 35 standard French + 9 additional in Swiss French), a matrix of all distances or all proximities has 8,649 (93 × 93) entries; there are two symmetric triangular halves. We have used the agglomerative nesting (AGNES) algorithm of Struyf, Hubert, and Rousseeuw (1996, 1997; Kaufman & Rousseeuw, 1990), which works as follows. In the beginning, each phoneme is a small cluster, just by itself. At each subsequent step, the two nearest clusters are found and are joined to form a progressively larger cluster; the distance between two clusters is, at all times, the (unweighted) group average of the distances between the phonemes in one cluster and the phonemes in the other cluster. At the end, a single large cluster remains that contains all phonemes. The succession of clustering steps gives rise to a hierarchy of clusters, which can be graphically

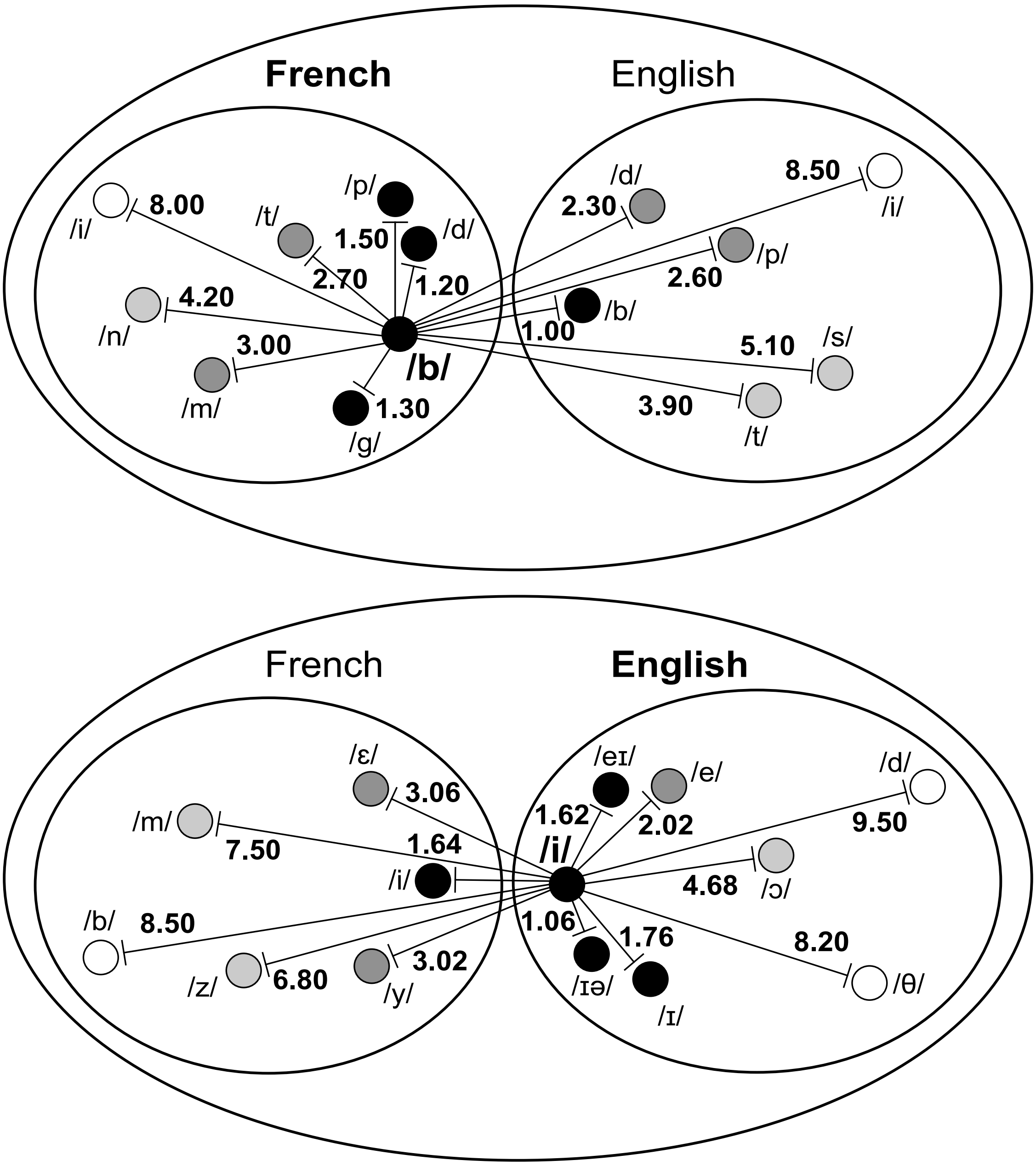


Figure 5. Distances from French /b/ (at the top) and from English /i/ (at the bottom) to a select number of other phonemes in the two languages.

represented in a tree-like structure, called dendrogram. The result for the whole bilingual matrix is a rather baffling, very large dendrogram. We therefore have created two monolingual dendrograms, one for English and one for Swiss French (they appear in Figure 6, on the next page), as well as two smaller bilingual/bidialectal dendrograms, separately for consonants and separately for vowels, glides, and diphthongs (they are shown in Figure 7, on the page after; E, F, and SF stand for English, standard French, and Swiss French).

In these figures, the phonemes are the leaves, that is, the end points at the very left, of the trees; close phonemes are grouped together first (i.e. more to the left in the trees), and groups of phonemes that are near to one another are combined later (i.e. more to the right in the trees). The left-to-right position at which two phonemes are connected corresponds to their distance ρ, and the position at which two groups come together is the average distance between members of the two groups. In both the monolingual dendrograms, we notice, in the upper part of the trees, the manifestation of separate clusters of plosives, fricatives, nasals, and liquids (and affricates for English), all merged together to form a main cluster of consonants; in the lower part of the trees, we discern a small cluster of glides, and a large cluster of vowels (and diphthongs in English), formed by a combination of subclusters of higher and lower vowels. Incidentally, the English allophones (e.g. [pʰ] and [p]) are joined next to each other, and the Swiss French long vowels (e.g. /uː/) are grouped with their short counterparts (/u/). With the aid of the bilingual dendrograms, we can study distance relations across English and French, and we can find interlingual phoneme neighbors that are remote, such as English /ɹ/ and French /ʁ/, or English /ɪ/ and French /i/, or close, like English and French /b/, /g/, or /n/. The phonemes /f, v, s, z, ʃ, ʒ, m, ŋ, j, w/, which have the same feature values across the two languages, have a distance of 0 and therefore share a single leaf per phoneme.

To complete our analysis, we computed the agglomerative coefficient, which measures the overall quality of the clustering structure found and grows with the number *N* of items in the tree (Struyf et al., 1996, 1997). It was high, which is a sign of a good clustering structure: 0.86 for English (*N* = 49), 0.81 for standard French[7] (*N* = 35), 0.86 for Swiss French (*N* = 44), 0.89 for consonants (*N* = 44), 0.83 for vowels, diphthongs, and glides (*N* = 49), and, finally, 0.91 for the whole matrix (*N* = 93).

7 The standard French dendrogram is not shown but similar to the Swiss French one.

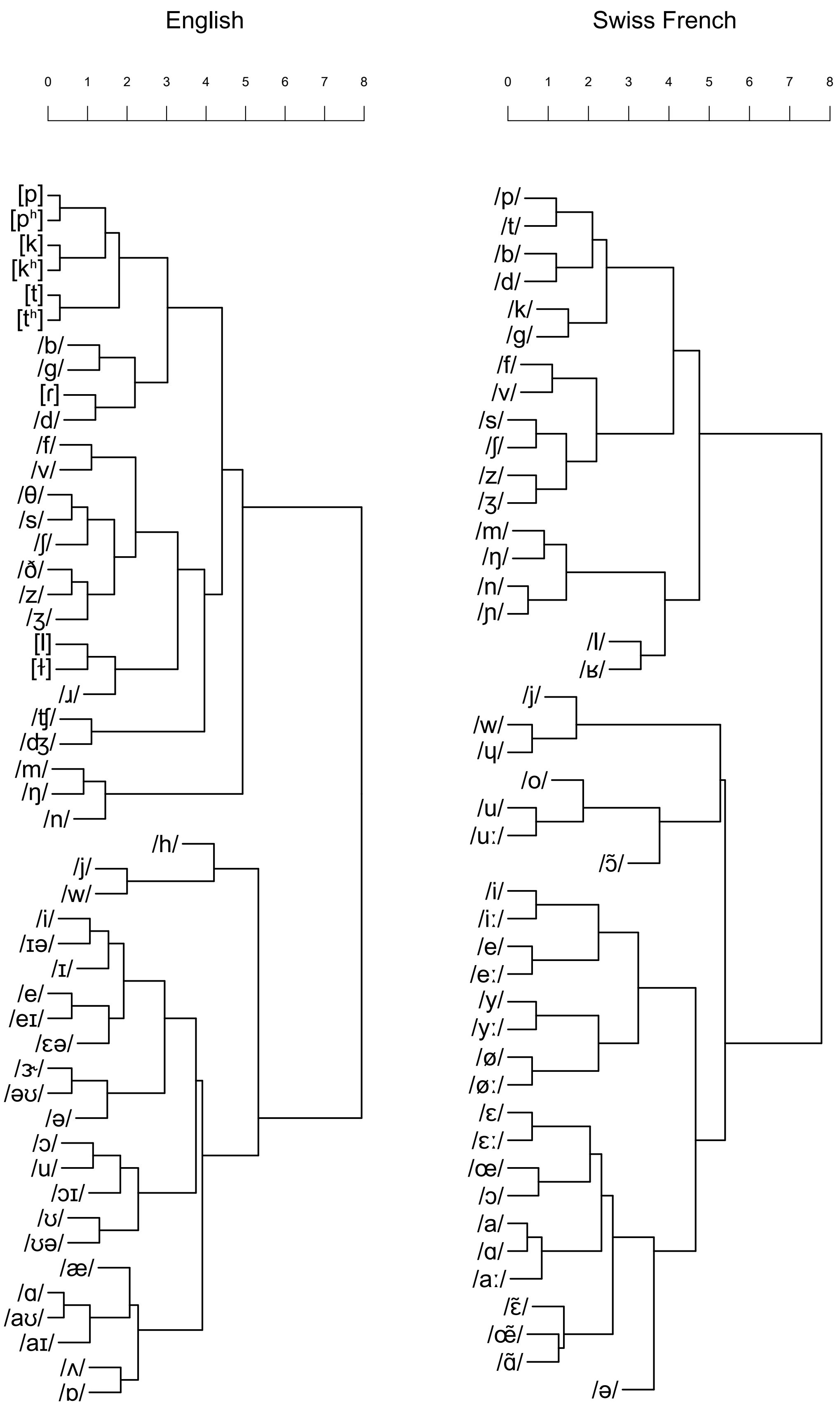


Figure 6. Hierarchical clustering of monolingual phoneme distances.

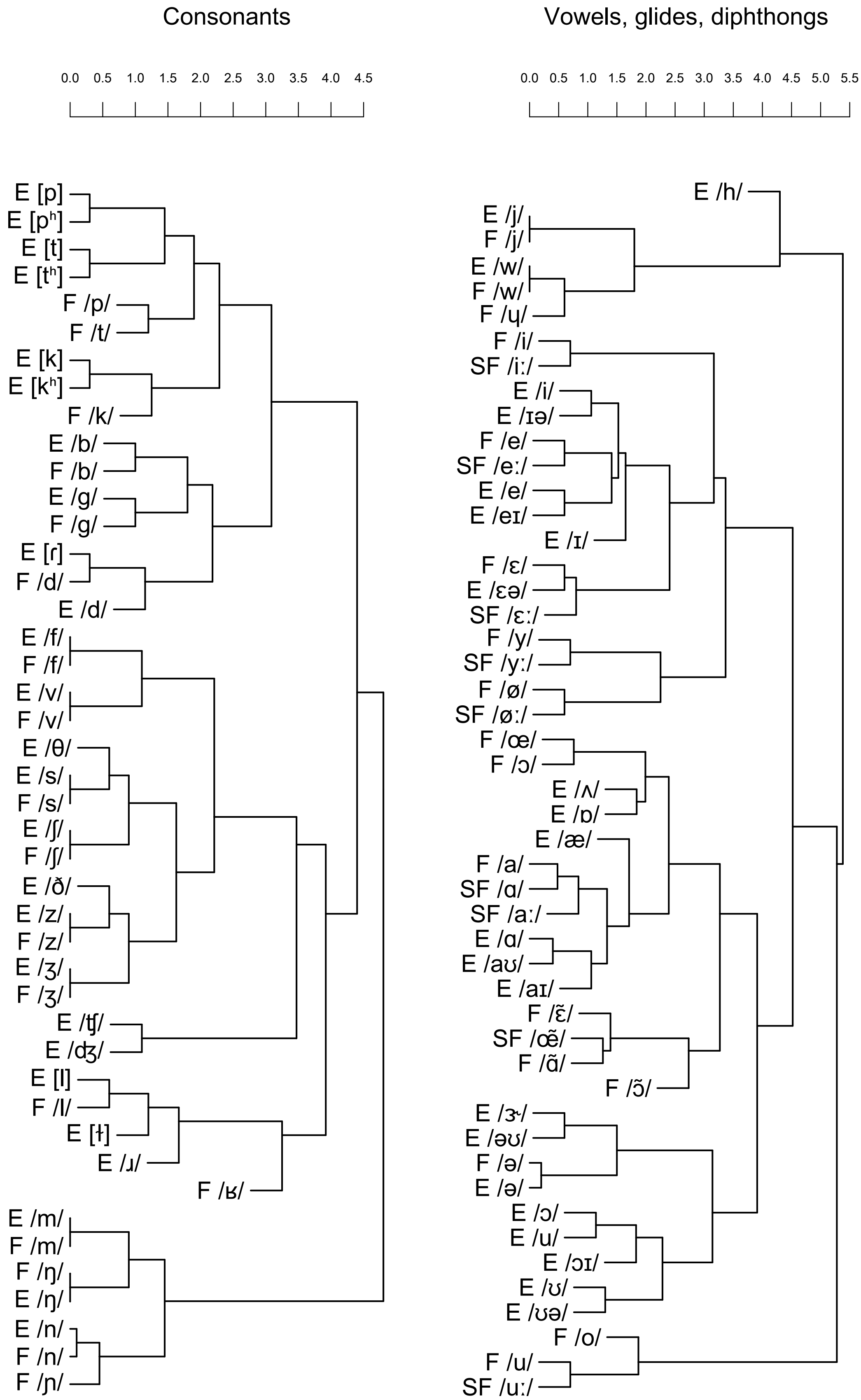


Figure 7. Hierarchical clustering of bilingual phoneme distances.

We now move up to the level of words and the lexicons, which we discuss in individual sections by model, first for BIMOLA and later for FN5.

Words for BIMOLA

For the bilingual model BIMOLA, we have created an English–French bilingual lexicon of 8,696 words, 4,348 for each of the two languages.[8] They are all verbs, because the experiment by Grosjean (1988), which is the principal study we will employ to evaluate BIMOLA, used the sentence frame “Il faudrait qu’on…” (“We should…”), which went on with a verb. The participants in this study were aware that they would hear verbs and had to propose verbs, either English or French. When a verb was in French, it was in the third person singular subjunctive (e.g. “Il faudrait qu’on choisisse” in the case of “choisir” (“choose”)). When a verb was in English, resulting in a mixed-language utterance (e.g. “Il faudrait qu’on skip”), it was in its base form. The verbs in the BIMOLA lexicon are listed in precisely these forms (e.g. “choisisse” in French and “skip” in English) and will be used in isolation, that is, as single words. Restricting the BIMOLA lexicon to verbs can be seen as an implicit top-down constraint in the model’s interactive activation network; like the participants in Grosjean’s study, BIMOLA will only hear verbs and will have to propose verbs. BIMOLA’s coverage of verbs is quite broad; and FN5 will take charge of some other lexical categories (nouns, determiners, and adjectives, albeit only in French) to round off the picture.

As we will describe below, we have first built two monolingual lexicons, one for English and one for French, taking very similar steps (partly automated, partly manual). We started by extracting linguistic information from an available lexical database; we continued by converting, checking, and correcting this raw data and by enriching it further; and we finished by removing as well as adding some words. Once the two monolingual lexicons were ready, we equalized the number of words in the two lexicons and normalized the frequencies of words across the languages (see further down for details). This allowed us to combine them into a single English–French lexicon. This bilingual lexicon is conceptually one large system comprising all the words of the bilingual person; at the same

---

[8] Considering the small-sized lexicons of some previous models (notably TRACE, McClelland & Elman, 1986), which tended to be hard to be scaled up to more realistic, larger lexicons (cf. Frauenfelder, 1996), we decided to construct the lexicon for BIMOLA with several thousands of words right from the beginning.

time, it encompasses two subsystems (i.e. the monolingual lexicons), which contain, respectively, the English and the French words (see the superset and the two subsets at the word level in BIMOLA's architecture; Figure 4, p. 17).

A short excerpt of the BIMOLA lexicon, showing the first 15 English and the first 15 French verbs of the letter F, is presented as Table 3. Both parts of the lexicon, English and French, contain the same linguistic information for each word listed: (a) the spelling; (b) the pronunciation (phonetic transcriptions are shown in IPA notation on screen and in this text, but are stored as equivalent one-byte character codes on file; for these, see the last column in Appendix A); (c) a number between 0 (lowest) and 1 (highest frequency) representing the word's frequency of occurrence; and (d) the uniqueness point. The values for the last-mentioned, an important variable of lexical statistics, depend on a given lexicon's entire content; indeed, we calculated these values, separately over all the BIMOLA lexicon's English words and over all its French words, using our own implemented algorithm.[9]

The uniqueness point (UP) expresses where within the beginning-to-end sequence of its phonemes a word starts to differ from all the other words of the same language and becomes unique (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1984). As a first example, the UP of the word "look" (/lʊk/) is on the /ʊ/, since no other English verb in the BIMOLA lexicon begins with the sequence of phonemes /lʊ…/. As /ʊ/ is the second of the three phonemes of "look", we have a value in percentage of word length of 66.7% (or 2⁄3). As a second example, the UP of "put" (/pʰʊt/) is on the /t/ (UP = 100% or 3⁄3) due to the fact that both "pull" and "push" differ from "put" only on that very last phoneme. Finally, some words do not even become unique at their end, for example "win" (/wɪn/), which is embedded in "winnow", "wince", "whinny", etc. (i.e. the phoneme sequence /wɪn/ can be the whole word "win" or the beginning of one of those other words). The UP of "win" depends on the phonemes following the /n/ but is surely larger than 100%. It is stored in the lexicon as 133.3% (or 4⁄3). As for examples from BIMOLA's French verb lexicon, "oublie" becomes unique on the /b/ (UP = 50% = 2⁄4), "amuse" on the /y/ (UP = 75% = 3⁄4), "garde" on the /d/ (UP = 100% = 4⁄4), and "reste" does not become unique until after word offset (UP = 125% = 5⁄4, the last a substitute for a UP > 100%).

---

[9] This variable, which will be used in the evaluations, is for information purposes only and has no effect on BIMOLA's functioning, even if we directly incorporated it into BIMOLA's lexicon. For FN5, we chose to store such variables in a separate lexical statistics database, which is the better solution.

Table 3.
Extract from BIMOLA's primary lexicon file "BimolaVerbs.Lex"

Pronunciation	*Spelling**	*Uniqueness point*	*Frequency*
	English examples		
/fæbɹɪkeɪt/	"fabricate"	37.5	0.104
/fəsɪlɪteɪt/	"facilitate"	44.4	0.208
/feɪd/	"fade"	100.0	0.301
/fæg/	"fag"	100.0	0.000
/feɪɫ/	"fail"	100.0	0.469
/feɪnt/	"faint(H)"	100.0	0.000
/feɪk/	"fake"	100.0	0.104
/fɔɫ/	"fall"	133.3	0.518
/fɔɫsɪfaɪ/	"falsify"	57.1	0.104
/fɔɫtə/	"falter"	80.0	0.169
/fəmɪljəɹaɪz/	"familiarize"	30.0	0.000
/fæmɪʃ/	"famish"	60.0	0.000
/fæn/	"fan"	133.3	0.243
/fænsɪ/	"fancy"	80.0	0.184
/fɛə/	"fare"	100.0	0.066
	French examples		
/fabʁik/	"fabrique"	50.0	0.304
/faʃ/	"fâche"	100.0	0.273
/fasilit/	"facilite"	71.4	0.227
/fasɔn/	"façonne"	80.0	0.158
/faktyʁ/	"facture"	50.0	0.000
/fagɔt/	"fagote"	100.0	0.000
/feblis/	"faiblisse"	50.0	0.158
/faj/	"faille(H-oir)"	133.3	0.709
/feneɑ̃t/	"fainéante"	50.0	0.000
/fəzɑ̃d/	"faisande"	40.0	0.000
/falsifi/	"falsifie"	42.9	0.000
/familjaʁiz/	"familiarise"	30.0	0.000
/fan/	"fane"	100.0	0.108
/faʁsis/	"farcisse"	66.7	0.000
/faʁd/	"farde"	100.0	0.068

* "H" marks homophones: English "faint" (homophone "feint"); French "faille" from "falloir" (homophone "faille" from "faillir").

Preparing an English lexicon. The OXFORD psycholinguistic database (Quinlan, 1992), an updated and corrected version of the MRC psycholinguistic database (Coltheart, 1981), made accessible to Macintosh computers, contains various linguistic information on a total of 98,538 English words. We extracted from OXFORD the spelling, pronunciation, and lexical category of all its words, and found that 38,292 words are listed with their pronunciation, 6,289 of which are verbs. These verbs were used as our starting point for BIMOLA's English lexicon and were processed as follows.

We converted the OXFORD spelling from uppercase to lowercase, and checked and corrected it where faulty (e.g. "lenghthen" [*sic*] [*recte* "lengthen"] or "sheperd" [*sic*] [*recte* "shepherd"]). The OXFORD pronunciation is, of course, in British English and also includes syllabification and stress patterns; it uses a (computationally inefficient) set of both one- and two-character codes, which we easily translated to our own set of only one-character codes. Again, we caught miscellaneous errors (such as /fIlt/ [*sic*] for "filter"). Since OXFORD does not differentiate the allophonic variants we have in our list of English phonemes, we applied two correction rules (in agreement with Gimson, 1989). The first one states: the fortis plosives /p, t, k/ are to be aspirated if placed at the beginning of a syllable with primary or secondary stress (like in "push", "play", "appear", and "palisade"), but unaspirated elsewhere (like in "spend", "hope", and "happen"). The second rule states: /l/ is to be clear if it precedes a vowel, diphthong, or /j/ (e.g. in "look", "blow", and "value"), but dark otherwise (e.g. in "help" and "feel"). Once these correction rules were carried out, we discarded the information on syllabification and stress patterns as it was of no further use to BIMOLA.

As regards word frequency, we could not use the values from Kučera and Francis (1967) that are included in OXFORD: they were based on the yet untagged Brown Corpus and do not distinguish between the parts of speech (or lexical categories). This is a problem for us because many English verbs are homographic with nouns and have different frequencies. For example, "table" has the same high value in Kučera and Francis (1967), for verb and for noun, although it is much more often found as a noun than as a verb; by contrast, the verb "to find" is more frequent than the noun "find". In a second analysis, based on the Brown Corpus now tagged for part of speech, Francis and Kučera (1982) counted words separately for every lexical category; these are frequency values we could use.[10] Since we did not have Francis and Kučera (1982) in electronic

[10] At the time we carried out this task, any more recent English word frequency counts (e.g. Brysbaert & New, 2009; Pastizzo & Carbone, 2007) were not available yet.

form, but only as a book, the value for each of our English verbs had to be read and entered into the computer by hand. When a word was missing in the book, a value of 0 was entered.

For a number of reasons, we removed words from or added words to our lexicon. First, we searched for pairs of homographs among the verbs and found duplicate entries (one of which was discarded), the same words with alternative pronunciations (only the better known of which was retained; e.g. /bleɪzn/ rather than /blæzən/ for “blazon”), and a few heteronyms (both were kept and specially marked; e.g. “lead(/i/)” and “lead(/e/)”). Second, we eliminated spelling variants (e.g. for “align/aline”, “scallop/scollop”, and “curtsy/curtsey”, the second variants had to go). Third, we deleted all inflected forms (e.g. “am”, “took”, “risen”, etc.) and contracted forms (e.g. “I’m”, “it’s”, “we’re”, etc.). Fourth, in order to remove very rare items, we asked a native speaker of British English to examine each of our words and to judge whether it really was a verb or not and, when in doubt, to check in the *Merriam-Webster’s Collegiate Dictionary* (1990) if it was perhaps acceptable in American English. About 300 entries (such as “hovel”, “frustify”, and “obvert”) were neither known as verbs to the native speaker nor listed as verbs in Webster’s, and were therefore deleted. Fifth, we identified auxiliaries (e.g. “be”, “can”, “have”, etc.) and put them away into a secondary file. The four forms used both as auxiliaries and as regular verbs (“can”, “dare”, “need”, and “will”) were listed, with appropriate frequency values, both in the main lexicon and the secondary file. We also moved some archaic verbs (e.g. “beseem”) into a file apart. (The secondary files are not used in our simulations but they can be loaded optionally; see Appendix B1 for a list of them.) Sixth, we tracked down all homophones (“break” and “brake”, for example, are both pronounced /bɹeɪk/; and “paw”, “pore”, and “pour” are all three pronounced /pʰɔ/, in British English). We marked them with an “H”, ranked them by frequency, retained in the main lexicon the most frequent item for each pronunciation (i.e. “break(H)”, “pour(H)”), and put aside the less frequent one (or ones) into another file. The idea was to make sure that never more than one word corresponds to any sequence of phonemes; the one word, as a representative for all the words with the same pronunciation, is bound to become unique and to be isolated at some point in time.[11] Seventh, we verified whether all the verbs that were proposed at least

[11] This precaution, taken for BIMOLA only, is not really necessary. In FN5, we left the homophones in the lexicon but instead redefined the uniqueness point (to disregard homophones) and also adapted the process of word isolation (i.e. a word is isolated even in the presence of more frequent and more highly activated homophones).

twice in the candidate responses of Grosjean's (1988) study were part of our lexicon; we added the eight words that were absent ("bill", "fit", "full", "peek", "pig", "scape", "snack", and "spot"), with frequency values again from Francis and Kučera (1982). We noted that a few other verbs of higher frequency are missing in OXFORD, and hence in BIMOLA ("clear", "dry", "freeze", "observe", "place", "urge", etc.). Although we could have justified, from case to case, to insert such words in our lexicon, the process would have been quite arbitrary (should we also include "glaze" and "panic", two other absent verbs but of lower frequency?) and potentially infinite (there is always another verb we could throw in, too). Therefore, we did not to add any more words to our lexicon. (Users of BIMOLA can add manually any word they care for and run their own simulations with an enlarged lexicon if they want to.)

Preparing a French lexicon. For BIMOLA's French lexicon, we based ourselves on BRULEX (Content, Mousty, & Radeau, 1990), a lexical database from Brussels, Belgium, which provides linguistic information for 35,746 French words and has been widely used in psycholinguistic research on this language. The more recent LEXIQUE database (New, Pallier, Ferrand, & Matos, 2001; New, Pallier, Brysbaert, & Ferrand, 2004) was not known when we worked on the French words for BIMOLA, but it will be of use for word frequency in FN5. We extracted from BRULEX the spelling, pronunciation, lexical category, and frequency of all its words (the frequency is composed of two data fields, see below). We identified 4,334 verbs to start with.

In much the same way as for the English words, there were many tasks involved in processing further the French words. The BRULEX spelling had to be reformatted regarding letters with diacritics (é, à, ô, ï, ç, etc.). The BRULEX pronunciation had to be converted to our own set of phonetic characters. Next, all the instances of /ɑ/ and /œ̃/ were turned to /a/ and /ɛ̃/, because the BIMOLA lexicon is concerned with standard French (in the Swiss French lexicon of FN5, we will have /ɑ/ and /œ̃/). Word-final schwas, which occur in BRULEX for words like "mettre" ending in a consonant + "-re" but which are not usually pronounced, were eliminated. Word-initial single quotes were also deleted; they mark words that begin with an aspirated *h* and that do not experience liaison in connected speech (e.g. "haleter"); we have no use of this information for the recognition of single words in BIMOLA (but by contrast, we will employ it for the recognition of multiple words in FN5).

Concerning the word frequency, there are two data fields in BRULEX: the form frequency combines the lexical categories that a given orthographic form belongs to (e.g. “conseiller” as verb and as noun); the lexical frequency distinguishes between them (Content et al., 1990, p. 559f). While the latter is much preferable, we found that it is missing for some 63% of the homographs (as calculated on the entire BRULEX database). To complicate things further, when lexical frequency is missing, BRULEX lists the form frequency arbitrarily in one homographic entry and a flag in the other entry or entries. Consequently, we chose the lexical frequency whenever it was present; if absent, we used the form frequency if that was listed; we manually copied the form frequency from another lexical category when we only found the flag; and finally, we entered a value of 0 when both lexical and form frequencies were unavailable.

As mentioned before, we need the French verbs to be in the 3rd-person singular of the subjunctive (for “Il faudrait qu’on…”) but BRULEX gives them, of course, in the infinitive. We developed a program, from scratch, that generates this particular form from the infinitive form (both the spelling and pronunciation). It states, in a number of rule-like instructions, how the endings of the infinitive have to be reduced, modified, and replaced to create the subjunctive. It should be recalled that the subjunctive in French is not formed regularly (*Le nouveau Bescherelle*, 1966; Roller, 1979; Grevisse, 1990): there are four groups of verbs (ending in “-er”, “-ir”, “-oir”, and “-re”, respectively), and each group is further divided into subgroups that follow a certain pattern (e.g. “modeler: modèle” vs. “jeter: jette”, or “choisir: choisisse” vs. “tenir: tienne”). There are also exceptions (e.g. “aller: aille”, “mourir: meure”, “pouvoir: puisse”, etc.), there exist defective verbs that do not have a subjunctive form (13 words; e.g. “ravoir” and “gésir”; we eliminated them), as well as verbs with two non-homophonous alternative forms (23 words; e.g. “asseoir: asseye/assoie”, “balayer: balaye/balaie”, etc.; we duplicated these verbs). Once all subjunctive forms had been automatically generated, we asked a native speaker to verify their correctness (both spelling and pronunciation). After that, we stripped all words of their infinitive and sorted them by their subjunctive.

Like for our English lexicon, we had reasons to remove some words from our French lexicon and to add a few others. We cleared up identical entries and variant spellings (e.g. for “becquette/béquette” or “parafe/paraphe”, we kept the first but not the second spelling). Again, we sought out all the homophones; we found not only ordinary homophone pairs (e.g. /pɑ̃s/ for “pense” and “panse”) and triples (e.g. /sɛl/ for “scelle”, “selle”, and “cèle”), but also ten pairs of verbs that become homophonous and homographic just for the subjunctive form (e.g. “peindre/peigner: peigne”, pronounced /pɛɲ/). We labeled regular homophones with an “H” (e.g. “pense(H)”) and subjunctive-only homophones additionally with their infinitive ending ("peigne(H-re)" vs. "peigne(H-er)"). Basing ourselves on a ranking by word frequency, we decided which homophone to retain in the main lexicon (the most frequent one) and which to put into a file apart. Then, also for our French lexicon, we made sure that all the verbs proposed more than once as candidates in Grosjean's (1988) study were indeed in the lexicon. Five words were not and had to be added ("ligne", "liste", "note", "se quitte", and "se tire"). Incidentally, "note" was used as a stimulus word in the Grosjean study; had we omitted it, we would not be able to test it in the BIMOLA evaluation. Since these words were missing in BRULEX, we did not immediately know their frequency and conducted a mini-study with 10 native speakers to obtain an estimate. With "se quitte" and "se tire" we had opened the door to reflexive verbs (*les verbes pronominaux*), and therefore we asked a native speaker to examine each of our verbs with respect to their use. For those classified as being never reflexive (e.g. "déjeune", "veuille") or occasionally reflexive (e.g. "(s')active" and "(se) charge"), no action was taken. But for the verbs classified as always reflexive (53 words; like "s'abstienne", "s'exclame", "se méfie", etc.), we added the reflexive versions to the lexicon (with the same frequency as "abstienne", "exclame", "méfie", etc.).

Putting the two monolingual lexicons together to form a bilingual lexicon. Once both the English and the French lexicons were ready, we made two final adjustments before putting them together into one bilingual lexicon. The first adjustment concerned the size of the lexicons. As we had as our objective the mental lexicon of an adult bilingual with more or less equal competence in the two languages (not a second language learner), the two lexicons had to have a similar number of words or, to make it simple, have the same number of words. But at that point, after our various additions and removals of words, the French lexicon had 4,348 entries whereas the English lexicon contained 5,453 entries, that is, 1,105 entries in excess. Deleting just any 1,105 English words without taking into account their frequency would expose us to the danger of removing accidentally some very frequent words, and taking away 1,105 words all of the frequency of 0 would alter the distribution of frequency of the remaining words. Instead, we proceeded as follows. First, we earmarked the candidate words of Grosjean (1988) so that they could not be removed. Next, we listed all the other words in our current English lexicon along with their Francis and Kučera (1982) frequency values (telling us how many times the words had appeared in the text corpus the authors analyzed). Then, we ran a random process that repeatedly selected one word and decreased its frequency value (its token number) by 1, until 1,105 words had received a frequency value of 0 (i.e. it was as if they had not appeared in the corpus at all).[12] Those are the English surplus words that we finally put aside into a secondary file, which will not be used for our simulations.

Our second adjustment aimed at normalizing the frequency values of all words in order to make them suitable for comparison across the two languages. Otherwise, BIMOLA would perhaps recognize French words more rapidly than English words, or vice versa, both of which we want to avoid. As is the custom, Francis and Kučera (1982) indicate the word frequency per 1 million words, but BRULEX (Content et al., 1990) gives the occurrence per 100 million words; we therefore first divided all the French frequency values by 100. We then replaced the absolute frequency values (or token occurrences) by their natural logarithm; this makes the values more practicable since they now grow linearly instead of exponentially. The idea to use log frequency values was already introduced in one of the earliest experiments on this variable (Howes & Solomon, 1951) and has, since McClelland and Rumelhart (1981, 1988), become common practice in computational psycholinguistics (see e.g. Coltheart et al., 2001; Dahan et al., 2001; Harm & Seidenberg, 1999; Luce et al., 2000; Seidenberg & McClelland, 1989; but cf. Plaut, McClelland, Seidenberg, & Patterson, 1996). In this regard, BIMOLA (and FN5, as we will see) is no exception. Finally, within each of the two languages, we normalized all words with regard to the word with the highest frequency: we divided all English log frequencies by the log frequency of “be”, and all French log frequencies by the log frequency of “soit” (subjunctive form of “être”). We thus obtained values ranging from 1 (most frequent) to 0 (very rare). It should be recalled that English auxiliaries are not part of BIMOLA’s primary lexicon but they can be loaded separately; the next most frequent English verb is “say” with a value of 0.749. For further examples, both English and French, we refer to Table 3.

After these two adjustments, the English and French lexicons were equal in size (4,348 words in each) and comparable as regards frequency values. As our last act, we therefore combined the two monolingual lexicons into a single English–French bilingual lexicon, containing 8,696 words. To conclude, we ended up with three primary and seven secondary lexicon files (as shown in Appendix B1). The primary lexicon files are the bilingual “BimolaVerbs.Lex” and

---

[12] When we have to do something at random (e.g. choose one item from a number of items), we use a random number generator we programmed ourselves (after Park & Miller, 1988; Reiser & Wirth, 1992). It provides floating-point and whole numbers, and tosses dice and coins, etc.

the unilingual “EnglishVerbs.Lex” and “FrenchVerbs.Lex”; these will be used in our BIMOLA evaluations. As secondary lexicon files, we find the various words that were set aside (i.e. homophonous verbs, auxiliaries, archaic verbs, and additional English verbs) as well as two large files that combine all words; the secondary lexicon files will not be employed in our evaluations but they can be loaded into BIMOLA as an option.

Box 1. To load a lexicon file into BIMOLA.

Open the “Lexicons” panel (via its top-left box), click “BimolaVerbs.Lex” to load the bilingual lexicon, or click “EnglishVerbs.Lex” or “FrenchVerbs.Lex” to load one of the monolingual lexicons. Click “See current” to view the lexicon content. One can also “Lookup”, “Add”, “Limit”, or “Remove” words, or “Remove all”. Finally, close the “Lexicons” panel (again via its top-left box).

Words for FN5

For the French-only FN5 model, we have prepared a substantial lexicon of 17,668 French words (i.e. four times as large as the French part of BIMOLA), 16,971 of which are nouns and 697 of which are determiners or prenominal adjectives, all in their singular form. They will be used in isolation, that is, one word at a time (e.g. “ami”, or “un”, or “petit”), and as sequences of two words: either determiner + noun (e.g. “un ami”) or prenominal adjective + noun (e.g. “petit ami”). Sequences longer than two words (e.g. “quel beau nouveau petit chien”) are, in principle, also possible with the FN5 lexicon but they will not be used or tested.

For the largest part, this lexicon has been based on BRULEX (Content et al., 1990), one more time. LEXIQUE (New et al., 2001, 2004) did not serve as a basis for the lexicon but played a part in determining the word frequency values (we will see below to what extent). We have built the FN5 lexicon in a similar way as we have set up the BIMOLA lexicon, that is, by taking words and their linguistic information from BRULEX and then reworking and enriching them greatly, both linguistically and computationally (we will not go into as much detail as we did above for the BIMOLA lexicon). As concerns linguistic checks and corrections, we received help from three French native speakers along the way, but did ourselves all the preparation, organization, and coordination work.

We extracted all the determiners, adjectives, and nouns of BRULEX. We decided on 18 determiners. Of 6,360 adjectives obtained from BRULEX, we sorted out a generous list of 679 prenominal adjectives, which are or can be placed before a noun (*épithètes antéposés*), in contrast to postnominal-only adjectives, which always go after the noun (*épithètes postposés*; Laenzlinger, 2005). Of 19,382 nouns extracted from BRULEX, we removed or put aside: plural-only nouns and forms (*N* = 330, e.g. “gens”, “agrumes”, “travaux”, etc.), nominalizations of adjectives that are already included (*N* = 328, e.g. “la jeune” and “le calme”), the names of the letters of the alphabet (*N* = 26, “le b”), as well as spelling alternatives and duplicate entries (*N* = 155). To avoid very rare or specialized terms (like “méhari”, “narthex”, “navaja”, “quasar”, “pélargonium”, etc.), our three French native speakers had to judge whether they knew the general sense of the 14,743 nouns with BRULEX frequency smaller than 500; for 1,609 nouns, two or three judged to not know; these nouns were eliminated. A handful of words of various interest (*N* = 37), including words that had served as stimuli in past psycholinguistic experiments in our laboratory (e.g. “ramequin” and “natel”), were added when they happened to be missing.

A small but characteristic extract from the FN5 lexicon, detailing twenty nouns (one line each) and five determiners/adjectives (partly over multiple lines), is presented as Table 4. The linguistic information that is present for each word consists of the following: (a) the spelling; (b) the pronunciation (as in BIMOLA’s lexicon, phonetic transcriptions are shown in the IPA on screen and are saved as 1-byte character codes on file; see the rightmost column in Appendix A); (c) the lexical category (either noun or determiner/adjective, this is stored by block of words); (d) the grammatical gender (masculine, feminine, occasionally both); (e) a numerical value representing the word’s frequency; and if applicable, (f) a set of 2, 4, or 8 preference values describing the usage of any schwa or schwas that the word contains (we will come back to it further down). Spelling, pronunciation, category, and gender have all been checked word by word, and corrected as needed, by our three French native speakers (each reviewed two thirds of the lexicon, working on computer using a verification tool that we had implemented just for that purpose). In addition, automated consistency checks were run on the entire data.

The whole lexicon has been prepared, and has been stored apart, for the two versions of French that the FN5 model covers, standard French (file name: “FrenchPtitami.Lex”) and Swiss French (file name: “SwissFrenchPtitami.Lex”). “Ptitami” comes from “petit ami” and alludes to the fact that these files comprise both determiners and adjectives, like “petit”, and nouns, like “ami”. Files that contain only determiners and adjectives, or only nouns, are also available, as are a few secondary lexicon files (see Appendix B2 for details). At the level of

Table 4.
Extract from FN5's primary lexicon file "FrenchPtitami.Lex"

Pronunciation	*Spelling*	*Gender*	*Frequency*	*Schwa usage*	
		Nouns			
/ekwatœʁ/	"équateur"	m	0.079		
/ekwasjɔ̃/	"équation"	f	0.156		
/ekɛʁ/	"équerre"	f	0.048		
/ekilibʁaʒ/	"équilibrage"	m	0.022		
/ekilibʁ/	"équilibre"	m	0.318		
/ekilibʁist/	"équilibriste"	m/f	0.029		
/ekinɔks/	"équinoxe"	m	0.051		
/ekipaʒ/	"équipage"	m	0.291		
/ekip/	"équipe"	f	0.402		
/ekipe/	"équipée"	f	0.076		
/ekip(ə)mɑ̃/	"équipement"	m	0.222	1.583	6.083
/ekipje/	"équipier"	m	0.114		
/ekipjɛʁ/	"équipière"	f	0.114		
/ekitasjɔ̃/	"équitation"	f	0.066		
/ekite/	"équité"	f	0.093		
/ekivalɑ̃s/	"équivalence"	f	0.094		
/ekivalɑ̃/	"équivalent"	m	0.192		
/ekivɔk/	"équivoque"	f	0.165		
/eʁabl/	"érable"	m	0.085		
/eʁadikasjɔ̃/	"éradication"	f	0.021		
		Determiners and adjectives			
/s(ə)gɔ̃/	"second"	m			
/s(ə)gɔ̃d/	"seconde"	f/f~			
/s(ə)gɔ̃t/	"second~"	m~	0.452	4.167	3.583
/sedɥizɑ̃/	"séduisant"	m			
/sedɥizɑ̃t/	"séduisante"	f/f~			
↑	"séduisant~"	m~	0.240		
/sɛzjɛm/	"seizième"	m/f/m~/f~	0.078		
/sɑ̃blabl/	"semblable"	m/f/m~/f~	0.331		
/sɑ̃pitɛʁnɛl/	"sempiternel"	m/m~			
↑	"sempiternelle"	f/f~	0.042		

Note. m: masculine, f: feminine; a tilde denotes a form used with liaison.

the pronunciation of words, three dialectal differences have been put into effect. Firstly, as mentioned, there is a conflation in standard French, but differentiation in Swiss French, of the /a, ɑ/ and /ɛ̃, œ̃/ contrasts (thus, e.g. “un” and “gâteau” have been transcribed /ɛ̃/ and /gato/ in the standard French but /œ̃/ and /gɑto/ in the Swiss French lexicon). Secondly, for Swiss French, long vowels have been inserted, where needed, at the end of over a thousand words (nouns like “vie”, “année”, “revue”, etc., and adjective forms like “vraie” or “jolie”, when pronounced in isolation); they are all feminine, by definition. (Long vowels in non-final positions, e.g. “reine”, have been left out to simplify things.) Lastly, as we will see below, the preference values for schwa usage also differ from standard French to Swiss French. The vocabulary itself was kept, on purpose, the same for the two versions of French, despite what a lexicographer of Swiss French (e.g. Thibault & Knecht, 1997) would be able to tell us. In this way, we will be able to pinpoint the source of FN5’s performance differences, between standard French and Swiss French, to these specific pronunciation differences. Unlike in the BIMOLA lexicon, we allowed all the homophones to remain in the FN5 lexicon (1,025 words for standard French, 605 words for Swiss French); /vɛʁ/ can be “verre”, “vers”, “vert”, “ver”, or “vair”: all of them will be activated in the model, and the most frequent one, “verre”, will be most highly activated.

Box 2. To load a lexicon file into FN5.

Open the “Lexicons” panel (via its top-left box), click either “FrenchPtitami.Lex” or “SwissFrenchPtitami.Lex”, then close the “Lexicons” panel again.

Words with pronunciation variants. While 85.7% of the words in the FN5 lexicon (i.e. the vast majority) have just one pronunciation, the remaining 14.3% of words are composed of multiple pronunciation variants: 2,352 words consist of two, and 173 words are made up of three or more (maximally eight) pronunciations, with the result that FN5’s 17,668-word lexicon contains a total of 20,523 pronunciations. (Breakdown by lexical category: the 16,971 nouns have 19,378 pronunciations, and the 697 adjectives and determiners have 1,145 pronunciations.)

All the pronunciation variants of a word are stored together and are considered a single representational entity, that is, represented as one word unit in the connectionist network, with just one level of activation. Pronunciation variants can contain preference values in order to distinguish more or less frequent variants of a given word and to lexically represent their respective strength. As schematized in Figure 8, there are three kinds of words that have pronunciation variants in FN5's current lexicon: (a) words that contain a schwa; (b) adjectives and determiners that take several word forms, due to gender inflection, to consonant liaison, or to both; and (c) words that can be contracted.

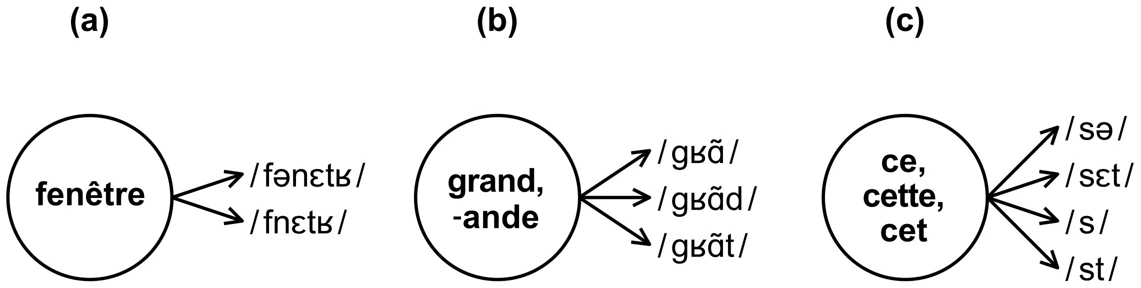


Figure 8. Schematic representation showing an example for each of three kinds of words modeled by FN5 using the approach of pronunciation variants.

Let us now examine each kind in turn and show how one converging method, namely the use of pronunciation variants, is capable of describing all these three important phenomena of the French language.

First, we recall that many French words containing a schwa (or "mute e"), such as the noun "f<u>e</u>nêtre", can be produced with the /ə/ either present, /fənɛtʁ/, or absent, /fnɛtʁ/; both the pronunciations occur and are perfectly legitimate. In linguistic theory, there are several competing views of how the deletion (or, as it is sometimes called, the absence and presence, or alternation) of schwa in a word should best be accounted for (see e.g. Delattre, 1951; F. Dell, 1985; Léon, 1966, 1996; Malécot, 1976, and many others). In the allomorphic approach of pronunciation variants we adopt here, "fenêtre" is specified within the lexicon as one word consisting of both these pronunciation variants (see Figure 8a). All the 2,190 words in the FN5 lexicon that have a schwa are represented in the same way. While schwa deletion is optional in the case of "fenêtre", it can shift all the way on a continuum from being mandatory via optional to prohibited (Hansen, 1994; Racine & Grosjean, 2005). For instance, in "contr<u>e</u>bande", schwa deletion is prohibited (i.e. the schwa must be present), whereas in "sci<u>e</u>rie", the deletion is mandatory (i.e. the schwa has to be absent). Apparently, the exact usage of the schwa depends on the word; it is a lexical property. For all the 2,190 schwa words, we therefore added preference values (one per pronunciation variant) into the FN5 lexicon. The values, provided to us and collected by Racine (2008), are the means of 12 natives' judgments, on an ordinal scale from 1 (lowest) to 7 (highest preference), obtained separately for standard French (in Nantes) and for Swiss French (in Neuchâtel). For standard French, the preference values for "fenêtre" (optional deletion) are 4.8, without deletion, and 3.4, with deletion; for "contrebande" (prohibited deletion), the values are 7.0 and 1.1; and for "scierie" (mandatory deletion), they are 1.0 and 7.0. There exist a few words that contain more than one schwa: 117 in the FN5 lexicon have two schwas (e.g. "revenu" and "développement"), and 3 even have three schwas (e.g. "ensevelissement"). We represent these words, by analogy, using a greater number of pronunciation variants and associated preference values: 4 (2^2) and 8 (2^3), respectively. (When the user types a target word into the word field, the transcription with the highest preference value automatically appears; the user can change it by manually adding or deleting a schwa.)

Second, adjectives and determiners in French typically possess several word forms, mainly because of gender inflection (masculine/feminine), but also as the result of the sandhi phenomenon of consonant liaison (see e.g. Delattre, 1947; F. Dell, 1985; Encrevé, 1983, 1988; Malécot, 1975; Morin & Kaye, 1982; Schane, 1968, etc., for discussions within theoretical linguistics). We were able to deal with both these phenomena (one usually considered morphological, the other phonological) by carrying on with the concept of multiple pronunciation variants. To begin with, while BRULEX has separate entries for feminine and masculine forms, we combined them in the FN5 lexicon into a single entry. The adjective "fort, forte", as an example, is represented as one word unit with two pronunciation variants: /fɔʁ/ (masculine) and /fɔʁt/ (feminine). Next, we added information on latent consonants of liaison in our adjectives and determiners. Words like "grand, -ande" (the example shown in Figure 8b) have stored three pronunciation variants: masculine "grand" /gʁɑ̃/, feminine "grande" /gʁɑ̃d/, and masculine with liaison, also written "grand" but pronounced /gʁɑ̃t/, with a linking /t/ at the end. We use a tilde (~) to show orthographically the presence of liaison, like in "grand~ amour", as opposed to "grand tambour", which is without liaison. When the user of the FN5 model types two words into the word field, the liaison is inserted automatically where appropriate. To prevent liaison, one can type a hash (#), like in "grand# ami", which would be transcribed to /gʁɑ̃ami/ (without liaison) and is uncommon.[13] For words like "tout, toute", the masculine liaison

[13] As the well-known rule states, appearance of liaison is triggered by the presence of a vowel, or a glide (/j, w, ɥ/), at the beginning of the next word (Encrevé, 1988; e.g. "grand~ opéra", "son~ hôtel", "un~ iota", "petit~ oiseau"). Exceptions to the rule are

form is pronounced exactly as the feminine form; these words have again two pronunciation variants: /tu/ (masculine without liaison) and /tut/ (masculine with liaison, or feminine); in the lexicon file, we use an arrow (↑) for this (cf. Table 4). We also had to deal with the following: words where the vowel varies between liaison form and feminine form (e.g. “entier~” /ɑ̃tjeʁ/ vs. “entière” /ɑ̃tjɛʁ/); words for which liaison is optional not obligatory (e.g. “profond(~) océan”); words with a third spelling form for the masculine with liaison (e.g. “bel” for “beau”); words that use a liaison form for both feminine and masculine (e.g. “mon~”, “deux~”, “cent(~)”[14]); words that are pronounced slightly differently when used in isolation or connected (e.g. feminine “gentille”), and so on. All these cases were covered indeed by the versatile approach of pronunciation variants. We suggest that it does not matter, at least for a psycholinguistic (not a linguistic) model, whether pronunciation variants are due to the existence of gender forms (masculine/ feminine), or whether they are the product of liaison: all are represented in the FN5 lexicon as pronunciation variants. For our approach, it is therefore neither a problem when some adjectives with gender and/or liaison forms also contain a schwa. Simply, the number of variants increases a little more. So, “petit, -ite” has a total of four variants: /pəti/, /pətit/, /pti/, /ptit/; and “premier, -ière” amounts to six variants (three adjective forms × two schwa variants).

Finally, we briefly turn to the third type of words that can make use of the pronunciation variants approach: words that can be contracted. It is well known that the French definite article “le, la” is shortened to “l’ ” before a vowel, as in “l’arbre”, “l’idée”, “l’automne”, etc. Even in writing, this elision is mandatory, and as a result, this form is omnipresent in French. What is perhaps less familiar is that a few more contracted word forms occur, as in “c’matin” (contracted from “ce matin”), “c’t après-midi” (from “cet après-midi”), “aut’ jour” (from “autre jour”), etc. (Arguably, these forms are not full words but clitics, see Matthews, 1991.)

---

two groups of words that block liaison, and for which there is a disjunction instead (Grevisse, 1993; Tseng, 2003): firstly, all words that start with an aspirated *h*, such as “héros”, “hasard”, “hauteur”, etc. (we marked them in the lexicon with an “H” at the onset of their transcription; a segment of silence, /−/, is inserted automatically when we use them in a simulation, e.g. /ɛ̃−eʁo/ for “un héros”); secondly, some of the words that begin with a glide, for example, “yaourt” (as in /ɛ̃jauʁt/) and “oui” (as in /ɛ̃wi/), words that we marked with an “X” in the lexicon. Most dictionaries (as *Le Petit Robert*, 1992) mark aspirated *h* (e.g. /’eʁo/) but few tell anything about words starting with a glide.

14 Adjectives only used before plural nouns are in a separate file.

Even though optional, these contractions are made quite often to facilitate the ease of rapid conversational speech (Malécot, 1976). Of course, they pertain to spoken not written French, but speech is the modality under examination here. Therefore, several contracted forms of determiners and prenominal adjectives (“l’ ”, “c’ ”, “c’t”, “c’te”, “not’ ”, “vot’ ”, “pauv’ ”, “aut’ ”, and “quat’ ”)[15] were included in the FN5 lexicon. They were implemented once again by using pronunciation variants. The demonstrative “ce, cette, cet”, for example, holds four variants: /sə/, /sɛt/, /s/, and /st/ (as visualized in Figure 8c).

Our approach of pronunciation variants is flexible and should be suitable for additional phonological/morphological phenomena in French as well as other languages. For example, would we intend to include plural forms of nouns and adjectives to our lexicon, we could attempt to represent them as variants of their singular forms. Both the storage of multiple variants and the existence of variant frequencies (like the ones we use for schwa words) have recently been verbally described, in a similar way to ours, for American English pronunciation variants pertaining to flaps and schwas (Connine & Pinnow, 2006; Connine, Ranbom, & Patterson, 2008). The FN5 model goes one step further and implements these notions computationally (albeit, of course, for French).

Word frequency values. Like BIMOLA, FN5 uses log frequency values, normalized to a range from 0 (very rare words) to 1 (very frequent words). The most frequent word of the French language, to which all words are normalized and which therefore itself has a frequency value of 1, is the verb “être” (“soit” in BIMOLA). It is not part of the FN5 lexicon as a verb but as a noun (frequency value of 0.438). Next follow the definite (“le, la, l’ ”) and indefinite articles (“un, une”), which are at frequency values of 0.989 and 0.941, respectively.

In contrast to BIMOLA, FN5’s frequency values are not based on a single frequency estimate per word, but on two: one derived from BRULEX and the other from LEXIQUE. These two lexical databases complement each other: BRULEX (Content et al., 1990), long-established and traditional, was based on formal, written language. By now, its frequency counts have arguably become somewhat dated; they originated from the *Trésor de la langue française* (Imbs, 1971) whose text corpus consisted of literary works published between 1919 and 1964. LEXIQUE (New et al., 2001, 2004) is a more recent lexical database for French, with frequency counts that are more up-to-date. In the version 3 (described by New, 2006), which we used, they were computed on a corpus of

---

15 “queq’ ” for “quelque” (as in “queq’ chose”) can be loaded from a separate file.

contemporary movies and television series subtitles: an innovative attempt to come a step closer to spoken (rather than written) language. To some, it would perhaps seem reasonable enough, had we simply taken the frequency values of LEXIQUE; but we do not agree. Combining the frequency information from two databases (which proved to demand a certain effort in correctly pairing it for each word) was worth it because it allowed us to use the two values to mutually correct one another. As a result, we obtained very reliable frequency values that we feel confident with. This is important because the frequency variable, as we will see later, has of course a strong effect on the model.

The procedure was the following. We paired, for every noun of the FN5 lexicon, the normalized log frequency value that we had derived from BRULEX (in just the same way as we have described above for the BIMOLA lexicon) with a normalized log frequency value that we were able to work out from LEXIQUE. A scatterplot, in which each point stands for one noun, is presented in Figure 9. As can be seen, the bulk of the words are packed in the (therefore very dark) lower left region of the figure. Moreover, the distribution of word frequencies of both lexical databases is in accordance with Zipf's law (G. Miller, 1957; G. Miller & Newman, 1958; Bard & Shillcock, 1993): whereas words of lower frequency (values close to 0) abound, there are fewer words, the higher the frequency (values away from 0).[16] We first attempted to model the relationship between the two variables, LEXIQUE vs. BRULEX, by means of simple linear regression analysis (method of least squares). However, the regression line is seriously biased by the presence of outliers, as can be seen in the figure (and as is often the case in linear regression); hence it was of no use. Instead, we identified the line that has a slope of 1 and an intercept of −0.047 (the mean of the difference between LEXIQUE- and BRULEX-derived log frequencies); this line traverses the points neatly. We added parallel lines of upper and lower limits of ±0.287 (3 standard deviations of the difference of log frequencies), and we thus cut the plot into three areas, A, B, and C, as is marked in the figure. Nouns in area A (N = 95) have a much higher frequency in LEXIQUE than in BRULEX. Most of these words are very informal, casual, or argot (e.g. "mec", "nana", "boulot", and "flic", to name but a few of the less offensive ones); some others (like "vampire", "assassinat", "shérif", etc.) seem to prevail in frequency because of predominant movie themes in LEXIQUE's cinematic corpus. To neutralize such undesirable tendencies in LEXIQUE, we gave to these outliers the (lower) BRULEX-derived frequency value. Nouns in area C (N = 170) possess, in contrast, a much higher

[16] We confirmed this with the help of histograms (not shown here).

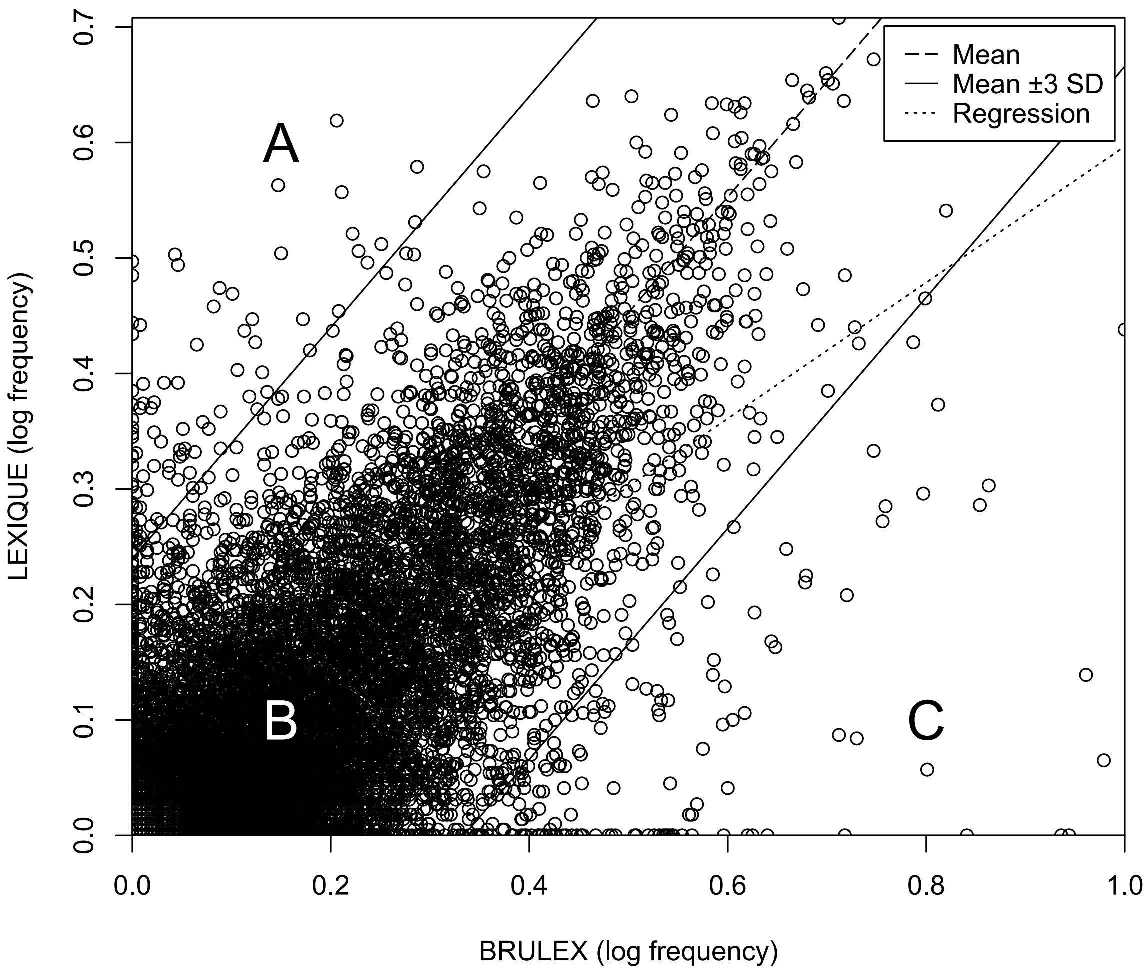


Figure 9. LEXIQUE-derived vs. BRULEX-derived normalized log frequency for all nouns in the FN5 lexicon.

frequency in BRULEX than in LEXIQUE. These nouns either come from topics one seldom talks about everyday (e.g. "métaphysique", "religieux", etc.), or else they have very frequent counterparts in another lexical category (e.g. "je", "si", "mais", "avoir", "vouloir", "aujourd'hui", "oui", "être", "y", "bien", etc.), from which the nouns in BRULEX had inherited the high frequency value (wrongly so, since "je" and the rest are hardly ever used as nouns). In order to correct this lack of differentiation in BRULEX, we assigned to those outliers the (lower) LEXIQUE-derived frequency value. Nouns in area B (i.e. the remainder and obviously the vast majority of words) received the arithmetic mean of the BRULEX-derived and the LEXIQUE-derived frequency values. Regarding the adjectives and determiners in the FN5 lexicon (they are not shown in the figure), we proceeded in just the same way and regulated their frequency by the nouns' boundary lines.

CHAPTER 4. GENERAL MECHANISMS

Among FN5's and BIMOLA's various internal mechanisms, the ones common to both models are described in this chapter; those specific only to FN5 or only to BIMOLA will be found in the next chapter. The mechanisms can be regulated by some parameters. They are listed in Table 5, in the order they appear in the user interface, along with their exact values in FN5 and in BIMOLA (for one or the other, in the case of model-specific parameters). A parameter is either just a binary setting (indicating whether the mechanism is switched on or off) or it is a numerical setting (specifying the force, length, level, etc., of the mechanism). As a rule, parameters concerning activatory mechanisms have positive values, and parameters relating to inhibitory mechanisms have negative values, in both our models. For the more traditional parameters, which are familiar from previous localist connectionist models, like from TRACE (McClelland & Elman, 1986) and the interactive activation model (McClelland & Rumelhart, 1981, 1988), it will be enough to simply state the values in the table. We will, of course, justify and explain the new parameters we propose (i.e. those marked with an asterisk in the table), at their appropriate place in the text. We will make sure one can get a good feel for their meaning and influence onto our two models, and time and again, we will offer instructions how one can try them out by oneself (by turning a mechanism on or off, or altering a parameter value, and exploring the result;

Table 5.
The parameters of FN5 and BIMOLA

	Parameter	*Value in FN5*	*Value in BIMOLA*
Single units	Minimum activation level	−1	−1
	Maximum activation level	1	1
	Decay of features	0.04	0.04
	Decay of phonemes	0.02	0.04
	Decay of words	0	0
	*Power output of phonemes	2	2
	Power output of words	4	4
Model I/O	*Phoneme input delay	16[a]	16[a]
	*Phoneme input extent	16	16
	*Difference for isolation point	0.001	0
	*Global language activation of phonemes		0.1
	*Global language activation of words		0.9
Between levels	Feature to phoneme activation	0.05	0.05
	Feature to phoneme inhibition	−0.1	−0.1
	Phoneme to word activation	0.02	0.02
	*Absent phoneme to word inhibition	−0.85	−0.85
	*Preceded words activation	0.35	
	*Words without continuation inhibition	Yes	
	Word to phoneme activation		0.01
	*Absent word to phoneme inhibition		0
Within levels	Lateral inhibition between phonemes	−0.35	−0.35
	*Attenuation guest language phonemes		4
	Lateral inhibition between words	−0.15	−0.2
	*Attenuation ending variants	6.5	
Biases	*Word frequency	0.5	0.5
	*Word frequency in lateral inhibition	Yes	No
	*Schwa deletion	−2.5	
Context	*Number of words	Yes	
	*Prenominal word at beginning	Yes	

* New parameters

[a] For fast speech rate: 3

see e.g. Box 3). Finding the right parameter values took some trial and error, a procedure which is often employed and typical for localist connectionist models (e.g. McClelland & Elman, 1986; Dell, 1986; Jacobs, Rey, Ziegler, & Grainger, 1998; Coltheart et al., 2001). But our models are fairly robust to a change to a parameter value (it simply alters a little that mechanism's force), and hence a systematic search, or even mathematical optimization, of parameter values was not at all needed. Moreover, the two models use almost identical values (we will justify the very few exceptions), which lends them added plausibility. Of course, once established, all the parameter values have been kept constant throughout our simulations and evaluations (except for phoneme input delay, as indicated in the table and for reasons we will turn to later).

Box 3. To manipulate a parameter.

Open the "Parameters" panel (via its top-left box), click "Inspect current", scroll down to the parameter to manipulate. Change the value, or remove or add the check mark next to it to switch the mechanism off or on. Finally click "Apply".

For example, increase the value of "Phoneme to word activation" from 0.02 to 0.03 or decrease it to 0.01. Word activation curves will be larger (or smaller) in volume, but they will have the same overall shape. Words may reach higher (or lesser) activation levels, but their isolation point will be practically the same. By contrast, switch off one mechanism totally, for example, remove the check mark next to "Word frequency" and click "Apply". Low- and high-frequency words will now not show any difference at all. (To finally go back to our default parameter setting, click "FN5.Diss.Set" or "Bimola.Diss.Set" in the "Parameters" panel.)

We now start with the activation and inhibition of phonemes, we will then continue with the activation and inhibition of words, and we will finish with the isolation (i.e. a sort of recognition) of words.

Phoneme activation and inhibition

The user of the model types a single English or French word (for BIMOLA), one or two French words (in the case of FN5), into the word field. This is the target, that is, the word or the sequence of words to be recognized by the model. If the target is found in the lexicon, its transcription appears automatically as a series of phonemes. For each of these phonemes, a pattern of feature values is taken from our feature matrix and is sent into the model, one phoneme after the next.

These feature values characterize the phoneme in question in just the way we need. They allow to calculate the similarities and differences that this phoneme has with all the other phonemes (and trivially with itself), according to our metric space of phonemes and distance measure (see Equations 1–3, p. 25f). While the phoneme proximity (Eq. 3) determines the amount of phoneme activation, the phoneme distance (Eq. 2) determines the amount of phoneme inhibition. We can combine the two into one expression, the phoneme input:

$$\text{input}(x, T, c) = \left(\alpha \left[1 - \frac{\rho(x, T)}{\max\limits_{(x, y)} \rho(x, y)} \right] + \beta \frac{\rho(x, T)}{\max\limits_{(x, y)} \rho(x, y)} \right) (1 - \delta)^{c} \qquad (4)$$

in which x is a phoneme; T is the target phoneme being presented to the model; c is a non-negative integer related to processing cycles (see below); and α, β, δ are the parameters, respectively associated with: feature to phoneme activation, feature to phoneme inhibition, and decay of features. The closer a phoneme x is to the target phoneme T, the greater its activation and the smaller its inhibition; and the farther it is from T, the smaller its activation and the greater its inhibition. For the target phoneme itself (i.e. x = T), the activatory term in the equation is at its maximum (= α·1), and the inhibitory term equals 0 (= β·0). For the phonemes farthest apart, the activatory term is 0 (= α·0), and the inhibitory term becomes maximal (= β·1).

In addition to the regular phoneme inputs, there is a silence input that denotes a pause (transcribed /−/). It is used to indicate the end of the word (or, for FN5, the end of the sequence of words) to be recognized. For this purpose, it is routinely appended to whatever is to be presented to the models. So, the input for the word "slew" in BIMOLA is /slu−/ (i.e. three phoneme inputs followed by one silence input), and the input for "un~ ami" into FN5 is therefore /ɛ̃nami−/ (i.e. five phoneme inputs plus the silence input). To signal the end of a word or sequence can be important whenever the phoneme sequence also corresponds to the beginning of another word (i.e. for the words that have a UP > 100%).[17] Indeed, /slu/ is also the beginning of "sluice" /slus/, but by entering not an /s/ but a silence after the /u/, "slew" can win over "sluice". This depends of course on word frequency: "slew" is less frequent than "sluice". Take "slow", for which the input is /sləʊ−/; "slow" is more frequent than "slope" and thus wins over "slope"

[17] 855 (9.8%) of the words in BIMOLA's lexicon; 3,481 (17.0%) and 3,291 (16.0%) of the pronunciations in FN5's standard French and Swiss French lexicons.

already during the diphthong; the silence input at the end only reinforces it more. The silence input can also be employed, in FN5, as an explicit boundary marker between two words of a sequence. For the very unusual sequence “petit tamis” for example, we might want to place a silence between “petit” and “tamis” (the input becomes /pti−tami−/) so as to distinguish it from “petit~ ami” (for which the input is /ptitami−/ too). But, as we will see in the evaluations, FN5 rarely needs such markers because it can usually find out the word boundaries itself. Silence has its own processing unit on the phoneme level, as it is the case in TRACE (McClelland & Elman, 1986) and in PARSYN (Luce et al., 2000), but not so in ARTWORD (Grossberg & Myers, 2000). To silence, all the phonemes have a normalized distance of 1 and proximity of 0. Therefore, when silence is fed into our models, there is no activation but maximal inhibition (as per Equation 4), for all phoneme units except silence itself.

In order to represent in our models the unfolding activation of a series of phonemes, phoneme units exist for each phoneme position. That is, there is a set of phoneme units for the first position (it represents the phoneme activations and inhibitions for the first phoneme input of the series), there is an identical but distinct set of phonemes for the second position (representing activations and inhibitions for the second phoneme input of the series), and so on, until the last possible phoneme position. Depending on the lexicon loaded, the models adapt their phoneme level automatically to the correct phoneme units (English and/or French phonemes for BIMOLA, standard French or Swiss French ones for FN5) in as many phoneme positions as needed (15 for BIMOLA: the longest word in its lexicon is 14 phonemes long + 1 position for the closing silence; 50 for FN5: longest word in lexicon (18), shifted to the end of the longest possible sequence of a prenominal adjective (13) and noun (18) and final silence (1), 18+13+18+1). The longest word in BIMOLA’s lexicon is French “internationalise”; the longest adjective+noun sequence in FN5’s lexicon is “extraordinaire internationalisation”. They all can be processed. But, of course, we normally work on shorter words.

Hartley and Houghton (1996) used, in their model of short-term memory for nonwords, a template in the form of a ring (five slots and a syllable boundary that connects the fifth slot back to the first slot) to parse a continuous sequence of phonemes into syllables. If we replaced our current linear phoneme structure with a similarly built circular structure (but surely with more phoneme positions than Hartley & Houghton’s five), we could allow our models to parse endless sequences of phonemes into words. There has not been any need for it as yet, since we worked on single words and two-word sequences so far; but it will be important to keep it in mind, once one or the other model is extended to longer sequences or to sentences.[18] The way we represent time in space, utilizing multiple phoneme positions in our models, is very much influenced by TRACE, except that we do not retain TRACE's phoneme positions halfway through a phoneme. Unlike TRACE however, our models do not reduplicate the words over time: BIMOLA processes only single words (they all have the one same alignment position), and although FN5 processes multiple words presented sequentially, it does so by using a position processor, or PP (as we will see later, all words exist once and find themselves at their best alignment position).

Timing and scheduling of the phoneme inputs to our models involves two parameters. Phoneme input delay, the first, tells how many processing cycles (i.e. iterations through the connectionist network) elapse between the start of one phoneme input and that of the next in the sequence of phoneme inputs. Phoneme input extent, the second parameter, specifies the number of cycles during which each of the phonemes is being input to the model at full strength (i.e. with c = 0 in Equation 4's rightmost factor). Once this full strength phase has passed, input of the phoneme continues, but with less and less strength (the c then equals the number of cycles since the phoneme input extent has ended; as a consequence, the phoneme input decays to 0 as time progresses). That the phonemes arrive at periodic intervals and that all phonemes have the same extent are, of course, severe simplifications, but they are present in most previous computational models, including TRACE and Shortlist (Norris, 1994).[19] Phoneme input delay and phoneme input extent are both set to 16 cycles, in the models' normal mode of operation. The exact number of cycles (which is close, incidentally, to Shortlist's 15 cycles per segment) is not of particular relevance, but the fact that both parameters are set to the same number is important. This corresponds to a strictly sequential presentation mode, in which the phonemes enter the recognition process successively, phoneme by phoneme, from the beginning to the end of the input sequence (first phoneme, second phoneme, and so on, up to the last phoneme in the sequence). We will see immediately below that this is not a fixed characteristic of our models and that we can just alter one of the two scheduling parameters in order to vary it.

As soon as phoneme units become active, that is, once they have an

---

[18] In a brief remark on how to adapt TRACE in a similar way, Norris (1994, p. 195) suggested that the size of such a structure might be determined by memory span.

[19] In our feature matrix, [LONG] describes the relative length of phonemes; so, the models' input does contain knowledge on longer vs. shorter phonemes, but this information is static, and in numbers not in time, as one would obviously prefer.

activation level that is greater than 0 (which, for some phoneme units, happens in the very first cycle already), they compete with all the other phoneme units in the same position by means of conventional and mutual lateral inhibition. For smoothing purposes, output signals of phonemes are always multiplied by themselves (i.e. power output of phonemes = 2) before they are used in the lateral inhibition between phonemes (in an analogy to the squaring of output signals of words in TRACE; see McClelland & Elman, 1986, p. 20, fn. 1). Inactive phonemes, that is, those with an activation level less than 0, do not influence other phonemes, but they have an impact on words, as we will see.

Varying the phoneme input delay, and thus the speech rate. We have so far described a sequential, beginning-to-end presentation mode in which phonemes enter the recognition process one at a time. This is the normal mode and default setting for both our models, and it corresponds to the prevalent view in spoken word recognition's verbal theorizing (e.g. the revised Cohort model by Marslen-Wilson, 1987) and computational modeling (e.g. TRACE and Shortlist). An example of how this rationale can be made apparent within psycholinguistic experimentation is the gating paradigm (Grosjean, 1980, 1988, 1996), in which spoken words are presented to listeners by providing successive segments of increasing duration, starting with just the very beginning of and finishing with the whole of the acoustic stimulus. Another spoken word recognition research line has, however, questioned the importance of an entirely sequential presentation and has lead to several verbal models that favor a more simultaneous approach, where a word's (particularly salient) earlier and later segments play a part all at once (e.g. Christophe & Dupoux, 1996; Cutler & Norris, 1988; Grosjean & Gee, 1987; Wauquier-Gravelines, 1996, 1999). We think that a computational model should best try to incorporate both approaches (sequential and simultaneous) so that it can serve the purpose of comparing the two approaches and revealing their respective characteristics.[20]

It is by varying one of our parameters, phoneme input delay (i.e. the delay between the start of one phoneme input and that of the next), that we allow the user to, in principle, freely navigate along a continuum from sequential

[20] Also in the simultaneous approach, it is normally assumed that spoken words have some internal order (i.e. there is an first, second, third segment, etc., of a word, as implemented by our position-specific sets of phoneme units). Toscano, Anderson, and McMurray (2013) challenged even this, by finding evidence that spoken words may be activated by their anadromes (e.g. "tip" - "pit", "mug" - "gum", "leap" - "peel").

to simultaneous. The sequential presentation mode remains the models' default mode and is obtained when the phoneme input delay is set to 16 cycles, what we will call the normal speech rate. With a phoneme input delay that is greater than the 16 cycles, phonemes enter the recognition process more slowly and are spaced out somewhat (because the delay is larger than the extent); the speech rate is slower. If, however, the phoneme input delay is less than the 16 cycles, phonemes are sent into the model more quickly and overlap to some degree (the delay is then smaller than the extent); the speech rate is faster. Taken to its extreme, with a phoneme input delay of 0 cycle, all phonemes start together and coincide with each other, as it happens for letter strings in models of visual word recognition (e.g. Coltheart et al., 2001; Grainger & van Heuven, 2003; McClelland & Rumelhart, 1981). For the auditory modality, this entirely simultaneous mode seems plausible for very short words only; for longer words, it is merely hypothetical. A phoneme input delay of 3 cycles, a value that is still small but does not make us fall into the total simultaneity, appears reasonable; we therefore worked with it and contrasted it with the default setting of 16 cycles.

Figure 10 gets it right to the point. The phoneme input delay of 16 cycles (i.e. normal speech rate, sequential presentation mode) is depicted in the left

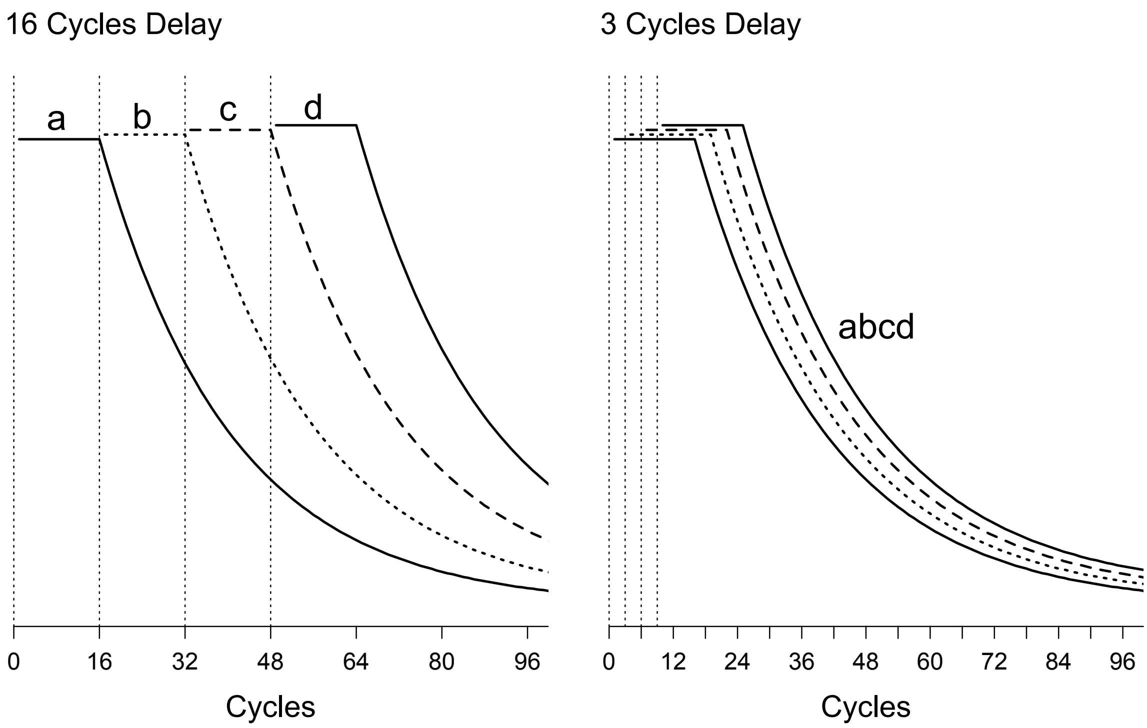


Figure 10. Strength of four phoneme inputs, a, b, c, d, when the phoneme input delay is 16 cycles (on the left) or 3 cycles (on the right).

panel of the figure, by means of a sequence of four phoneme inputs, abstractly labeled a, b, c, d. The first phoneme input begins at cycle 1 and is presented with full strength up to and including cycle 16 (since the phoneme input extent is 16 cycles, too). After that, the strength of this first phoneme input swiftly decays (the decay's shape being defined in Equation 4). The second phoneme input starts to the full at cycle 17, the third phoneme input at cycle 33, etc., and each phoneme input's strength similarly fades away in its time. We can see, in the diagram, that the level parts of the four curves (during the initial maxima) do not overlap; rather, they follow one another sequentially. In the right panel of the figure, the contrasting diagram for the phoneme input delay of 3 cycles is shown (i.e. fast speech rate, nearly simultaneous presentation mode). Here, the four phoneme inputs enter the process much more rapidly (note in the diagram the quick succession of the four dotted vertical lines): The first phoneme input commences at cycle 1, the second one at cycle 4, the third one at cycle 7, etc. The four curves retain the same shape and strength but they are now squeezed very close together. As the phoneme input extent still is 16 cycles, the curves' level parts (at their beginning) have now a considerable overlap. Nonetheless, they do not coincide; they are still in a sequence. Whereas we do not otherwise account for the challenging phenomenon of coarticulation of speech (see e.g. Elman & McClelland, 1988; Hardcastle & Hewlett, 1999; Nguyen, 2001), this overlapping of phonemes is a simple approximation.[21]

Word activation and inhibition

Let us move one level up and examine the general mechanisms for the words. Phoneme to word activation is obvious and of the commonplace style: A word is activated by the phonemes that it is made up of and that are currently active (i.e. have an activation level above 0). For example, French “tape” /tap/ is activated by /t/ in first phoneme position, /a/ in second phoneme position, and /p/ in third phoneme position, when indeed one or the other or all of these three phonemes are currently active in the respective positions on the phoneme level. In addition, “tape” is activated by the silence unit if that is active in the subsequent (i.e. the

[21] We found it to be fine to use the same strength of phoneme inputs whatever the phoneme input delay; we could have built a multiplicative gain control (similar to the one proposed by Grossberg et al., 1997), increasing vs. decreasing the strength of phoneme inputs for a faster vs. slower speech rate, but there was no need for it.

fourth) phoneme position. (As French “tape” happens to be a noun as well as a verb in the 3rd-person singular subjunctive, this example works both with FN5 and with BIMOLA.[22]) The phoneme to word inhibition in our models is, however, new and a little out of the ordinary, and therefore needs some more information. We have given already above (when we first introduced the two models, see p. 15) our reasons why bottom-up inhibition is present in our models and have clarified that it is gradual not categorical. Here we discuss which form of gradual inhibition we use. In fact, it seems to us that there exist two possible forms:

1. A word is inhibited by the phonemes that are active but that the word does not contain. For example, “tape” is inhibited, among other things, when /n/ is active in first position, or when /t/ is active in final position (i.e. in an incorrect position for “tape”).

2. A word is inhibited by the phonemes that it contains but that, at this moment in time, are inactive (i.e. have an activation level below 0). For example, “tape” is inhibited by /t/ in first, /a/ in second, and /p/ in third position, if these phonemes are not active on the phoneme level in those positions.

The first form of bottom-up inhibition is an inhibition of incompatibility, in the sense that phonemes inhibit the words with which they are incompatible. It is known since the interactive activation model (IAM). As one can see when one revisits the figure shown in McClelland and Rumelhart (1981, p. 380), it requires a fairly large number of inhibitory connections, because words are inhibited by all phonemes (i.e. all letters in the IAM) except the ones that they contain. Had we used this form of inhibition, “tape” would be inhibited, in the third phoneme position for example, by everything other than /p/. But such an approach poses a problem whenever several phonemes become active simultaneously for one and the same position (say, /p/, /t/ and /k/): these phonemes then tend to inhibit all words (even “tape”, to some degree), unless the inhibition is to be very weak. Indeed, in the IAM, letter to word inhibition was only about half as strong as letter to word activation (see McClelland & Rumelhart, 1981, p. 387).

The second form of bottom-up inhibition is an inhibition of absence, since an absent phoneme inhibits words that have this phoneme. It is rather new, as far as we know, and is the one we have employed in our models. It requires a much smaller number of inhibitory links, just as few as there are activatory links

---

[22] We counted 901 such words among BIMOLA’s French verbs and FN5’s adjectives and nouns. There would be more in English, but we lack an English noun lexicon.

from phonemes to words. In fact, the activatory and inhibitory connections run exactly from the same phonemes to the same words (e.g. from /t/, /a/, and p/ to "tape"), and they assume the corresponding roles of receiving and transmitting a state of presence or absence of phonemes to a (suitably converted) state of presence or absence of words. As a result, a word's bottom-up input is directly dependent on the activation levels of all of its constituent phonemes: For each of them, there is activation whenever the phoneme's activation level is greater than 0, and there is inhibition every time it is less than 0. In BIMOLA, "tape" is also inhibited by the silence unit when it is inactive in the subsequent phoneme position; in FN5, "tape" is likewise inhibited if no other words follow it (see the words without continuation inhibition, in the specific mechanisms). The exact amount of the bottom-up inhibition is determined by the absolute value of the phonemes' output signals and weighted by two parameters (one for activation, one for inhibition). Since this kind of inhibition is concentrated and specifically directed to just the right words, it can be increased without problem as much as one wishes. Indeed, in our models, absent phoneme to word inhibition is 42.5 times (i.e. considerably) stronger than phoneme to word activation (parameter values: −0.85 for inhibition vs. 0.02 for activation).

Box 4. To examine the role of "Absent phoneme to word inhibition".

In either BIMOLA or FN5, use the "Parameters" panel to switch off "Absent phoneme to word inhibition" (uncheck its check box and click "Apply"). Enter the French word "biche" (i.e. /biʃ−/ with a final silence). This word still reaches the highest activation level and is well recognized by either model, but it appears along with many mismatching candidates. In BIMOLA, there are, among others: French "bichonne" (because it is not inhibited anymore in spite of the silence), "bisse", "bise", and "biffe" (they are not inhibited anymore despite the /ʃ/), and English words like "beat", "base", "beset", etc. (use "Best 6" or "Best 30" to list the most activated candidates into the log). In FN5, they include (for the same reasons): "bichon", "bichonne", "bille", "bise", etc. Switch on "Absent phoneme to word inhibition" again, in either model. All the mismatching candidates will be promptly inhibited, and "biche" will be the only remaining word. This, of course, is the setting that we will use.

When word units become active, they compete with one another by means of mutual inhibition. As on to the phoneme level, output signals of words need to be smoothed before being used in the lateral inhibition between words; they are raised to their fourth power (i.e. power output of words = 4), due to the large number of words in both our models. This should be compared to TRACE, which used a much smaller number of words (a few hundreds), and for which a square sufficed, apparently, for the same purpose (McClelland & Elman, 1986, p. 20, fn. 1). In BIMOLA, the inhibition between the words (within each language, English and French) is equal for all words, because they all compete for the one same alignment position (parameter value: −0.2). In FN5, the inhibition between two words depends, as in TRACE, on the number of phoneme positions that these words have in common (parameter value: −0.15 per phoneme position). That is, the more the two words overlap, the higher their mutual inhibition (since overlapping words can usually not be part of the recognition result together). Two words that do not overlap do not inhibit one another (they can very well both contribute to the recognition result and might just form the right sequence of words).

Influence of word (and variant) frequency. Although TRACE and Shortlist did not deal with word frequency (Frauenfelder, 1996), there are several ways how one can account for this important variable in a localist connectionist model (see e.g. the extensions of TRACE by Dahan et al., 2001). In our models, word units have a resting level, an output signal, and a response strength, and each of the three could be the one affected by word frequency. The resting level of a word is the initial activation level from which a word starts. The output signal is what goes out of the word and influences other words, notably in the lateral inhibition between words. The response strength is the level that is displayed in a word graph and that determines whether the word surpasses the response strength of other words and, as we call it, is isolated (see further below for more information on word isolation).

In the standard way of modeling word frequency, the frequency is in the words' resting levels: the more frequent a word, the higher its resting level (e.g. we explored a range from −0.25 to +0.25). This gives high-frequency words a head start over low-frequency words, right at their very onset, and is an often employed method (see e.g. McClelland & Rumelhart, 1981; Stemberger, 1985; G. Dell, 1990; Dahan et al., 2001). We implemented it too but after we observed that the high-frequency words' growth was, on occasion, more than desirable and quite erratic, we did not continue to use this method. Besides, in BIMOLA, as we will see later, the resting levels of all the units (i.e. words and phonemes) are already dependent on the bilingual's language mode (resting level of 0 for base language units, and between −1 and 0 for guest language units, see p. 85). In FN5, all the units have a resting level of 0. In the two alternatives of modeling word frequency that we propose here, we use a frequency bias term, which is composed of the word's log frequency F and the current activation level L. Whenever L is greater than 0, the bias term is

$$(F - 0.5)\,(M - L)\,\gamma \tag{5}$$

with M being the maximum activation level; and when L is less than 0, it is

$$(F - 0.5)\,(L - m)\,\gamma \tag{6}$$

with m being the minimum activation level (cf. McClelland & Rumelhart, 1981). The modulation by the current activation level gives the word frequency a large effect as long as a word resides close to 0, but a small one once it approaches the maximum or minimum levels. The term, and hence the effect of frequency, is positive for words with a log frequency larger than 0.5, and it is negative for words with a log frequency smaller than 0.5 (of course, for the vast majority of words, 99.0% in FN5's lexicon and 97.5% in BIMOLA's lexicon, the latter is the case; see our discussion of the word frequency distribution in Figure 9, p. 48f). Setting the word frequency parameter γ to 0.5 (as we did) causes words to start with response strengths in the ±0.25 range, with which we had already worked using the standard approach.

Now, where and when is this frequency bias term inserted and added to? In our alternative approach A of modeling word frequency, it is in the words' output signals: the more frequent a word, the higher its output signal (which is the signal that influences other words), therefore the more the word is capable of suppressing competitor words by lateral inhibition. The response strength of a word is, in this approach, the same as its output signal; as a result, the more frequent a word, the quicker also its response strength exceeds those of other words, hence the sooner the word is isolated. To put it briefly, word frequency is modeled, in this approach, as an interactive activation bias. By contrast, in our alternative approach B, the words' output signal is not influenced by frequency, and therefore neither is the lateral inhibition between words; that is, frequent words do not inhibit competitors more than any other words do. The frequency is not applied any earlier than in the words' response strengths: the higher, the stronger (only for this method, the separate concept of a response strength is actually needed). Since the response strength is what determines when a word is isolated, high-frequency words are, also in that approach, isolated faster than are low-frequency words. We can say that word frequency is accounted for, in that approach, as a late operating response bias (cf. Connine, Titone, & Wang, 1993).

The approaches A (frequency already in the output signal and hence influencing the lateral inhibition) and B (frequency in the response strength and not yet in the lateral inhibition) can be compared by switching one parameter, named “Word frequency in lateral inhibition” (Yes gives approach A, No gives approach B). We used approach A (frequency in the lateral inhibition) in FN5, and approach B (frequency in the response strength) in BIMOLA. For BIMOLA, because it processes only single words, both approaches operate well enough and do not make much difference; with either approach, BIMOLA needs less bottom-up information for frequent words than for infrequent words. However, for FN5, we must opt for approach A (frequency in the lateral inhibition): word frequency must be allowed to have a bias effect on lateral inhibition in order to simulate the recognition of multiple, connected words. Let us take, for example, sequences with elided “l’ ” (like “l’acrobate”, “l’amateur”, “l’avocat”, “l’ancêtre”, “l’objet”, “l’orage”, “l’ouvrier”, “l’unité”, and so forth). For these, FN5’s preceded words activation mechanism is helpful (we will meet it later when we will present FN5’s specific mechanisms), but the influence of word frequency in lateral inhibition is even more important. Indeed, if frequent words inhibit more than less frequent ones do, there is a great deal of lateral inhibition from “l’ ” to the words that cover the initial (underlined) phonemes of the above sequences; that is, respectively, to “lac”, “lame”, “lave”, “lent”, “lobe”, “lord”, “louve”, and “lune”. In a similar way, the sequence “la sauterie” cannot result in “lasso” (because it is less frequent than “la”), “ta claque” does not lead to “tac” (less frequent than “ta”), neither does “ma jarre” produce “mage” (less frequent than “ma”), nor “du poids” give rise to “dupe” (less frequent than “du”), etc.[23]

Box 5. To examine the role of “Word frequency in lateral inhibition” in FN5.

First turn off “Context: prenominal word at beginning” (so that the first word of a sequence does not necessarily have to be a prenominal adjective or determiner but can be a noun). Test “la sauterie”, “ta claque”, “ma jarre”, and “du poids”, and toggle “Word frequency in lateral inhibition” between on and off. When it is off, FN5 will propose *“lasso tri”, *“tac lac”, *“mage art”, and *“dupe oie” (these contain exactly the same phonemes but less frequent words). Turn it back on: FN5 will produce the correct (i.e. more frequent) word sequences. Now also test a sequence with elided “l’ ”, say “l’unité”. Observe that “lune”, temporarily, is too good a candidate if “Word frequency in lateral inhibition” is off, but not if it is on. Notice that FN5 recognizes “l’unité” in either case, as there exist no word *“ité”.

[23] Ultimately, we plan to apply the more powerful approach A in all our models.

Apart from the word frequency bias, other biases that are intrinsic to the word, and expressed as a number, could be implemented in a similar manner, be it by replacing the F in the formulas (i.e. in lieu of word frequency) or be it by putting in a second term (i.e. additive to word frequency). Estimation values of the required variable (or variables) would have to be available for all the words of our lexicon. They would be transformed into the normalized 0 to 1 range that we use for word frequency (0.5: no bias, > 0.5: activatory bias, < 0.5: inhibitory bias), and be stored in the lexicon as well. Finally, the bias would need to be suitably weighted (by a well-chosen γ in the formulas).

Moreover, for words with multiple pronunciation variants, it is possible to consider, not only biases between words (like word frequency), but also biases within a word, that is, between the pronunciation variants of the same word. This is exactly what we did, in FN5, for the variable of schwa deletion, which obtained its own bias term, with quite a strong weight (γ = 2.5). As we have mentioned above, the FN5 lexicon contains preference values for the 2,190 words of the lexicon that can contain a schwa, encompassing words where the schwa deletion is mandatory, optional, or prohibited (Racine, 2008). The preferences (with or without schwa) are expressed as a separate value for each pronunciation variant, that is, they are a kind of variant frequency and can be used to implement a variant frequency bias. The (1 to 7) preference values in FN5's lexicon are converted into the (0 to 1) range of a bias effect as follows: All the pronunciation variants that have the schwa present are at 0.5 (no effect); pronunciation variants with the schwa absent are slightly above 0.5 if the schwa deletion is mandatory (small activatory effect), between 0.5 and 0 if the schwa deletion is optional (small inhibitory effect), and close to 0 if the schwa deletion is prohibited (large inhibitory effect). As we will see in the FN5 evaluations, this permits us to reproduce the results of Racine and Grosjean (2005).

Word isolation

In the experimental task of gating (Grosjean, 1980, 1988, 1996), participants listen to words that are presented in segments of increasing duration, from the beginning of a word to its end (different variants of the task exist). In the first presentation, they usually hear the word's initial part (e.g. the first 30 ms); in the second presentation, they hear a segment twice as long (e.g. the first 60 ms), and so forth, up until they hear the whole word in the last presentation (that is sometimes followed by additional gates containing subsequent words in order to disambiguate the context). After each presentation, participants are asked to quickly propose the word in question by writing it down or speaking it aloud (and to also rate their confidence). By examining these responses, the researcher establishes, for each participant, the isolation point of the word, that is, the point corresponding to the segment duration (in milliseconds or as a percentage of the word), within the sequence of presentations, in which the participant correctly identified the word, without thereafter modifying his or her response.

In Léwy et al. (2005), we suggested to adapt this measurement, for a usage in our models, in the following way. In every cycle of the simulation, the model (FN5 or BIMOLA) determines the response strengths of all the words, basing itself on several information sources: ascending activation from active phonemes, ascending inhibition from inactive phonemes, lateral competitive inhibition between words, and word frequency. We will see, in the next chapter, that FN5 (but not BIMOLA) also takes into account the words' current position and pronunciation variant, the status whether other words precede and follow them or not, and certain contextual knowledge (these factors do not play a role in the recognition of single words, as in BIMOLA). The contributing sources have partly an activatory and partly an inhibitory effect on the words' response strength. Words that resemble rather well the sequence of phonemes that is gradually being unfolded and that, in the case of FN5, also meet sufficiently the aforementioned other constraints (position, variant, preceding/following word, syntactic context), will accumulate more and more activation, but little inhibition, and thus will gain a progressively higher response strength, from one simulation cycle to the next. Several candidates will be entertained by the model in parallel and at different degrees (appropriately represented in the individual words' response strengths). But the target word, that is, the word to be recognized by the model, will be the one among the candidates to increase its response strength the fastest (at least in a normal course of simulation) because this word matches the incoming phoneme sequence the best or even perfectly.

The isolation point (IP), in BIMOLA and FN5, is therefore the moment in the simulation where the target word surpasses the response strength of all the other candidates and from which on it will retain this highest rank until the end of the simulation. (In FN5, to avoid situations of temporary commitment, the difference between target and other words, the target's homophones excepted, is required to exceed 0.001.) The IP measure is proposed in absolute terms of simulation cycles (= IP@cycle), as well as relatively, in percentage of the word's length including silence (= %IP), to neutralize the effect of length. For example, a word that is isolated at the cycle 24, and that consists of two phoneme inputs plus a silence input, of 16 cycles each, would have a %IP of 24 ⁄ [16 + 16 + 16] = 0.5 = 50.0%, as a result. The general equation is as follows:

$$\%IP = IP@cycle / [\text{phoneme input delay} \times \text{no. of phonemes} + \text{phoneme input extent}] \quad (7)$$

When a sequence of words is input to FN5, there are several target words and a separate IP for each; the IP is calculated from the beginning of each word.

There is no consensus, regarding Grosjean's paradigm of gating, on how much more time participants need, after a word's isolation, to reach its actual recognition; some specified a "recognition point", or a "total acceptance point", using the participants' confidence ratings (Grosjean, 1985; Tyler & Wessels, 1983; Walley, Michela, & Wood, 1995). Although we could define, for BIMOLA and FN5, an "RP" that occurs somewhat later than the IP (e.g. by demanding that the difference between target and other words exceeds a certain constant, say 0.1), we will assume for now that word isolation and word recognition are basically one and the same in our models. But, to keep this issue open, we will take care to always employ the phrases "the word is isolated", "word isolation", "isolation point", rather than "the word is recognized", "word recognition", and "recognition point". The IP, be it in terms of cycles or be it as a word-percentage, will be the one key measure used to relate our results with experimental results. That is, rather than choosing some measure for one simulation and a different measure for other simulations, we will have the IP measure followed through all our simulations, both for BIMOLA and FN5. The IP measure has the following properties. It expresses, depending on the particular word, one moment in time somewhere within the word or after it. It is calculated on the fly, that is, while the word progressively unfolds from its first to its last phoneme (whereas a measure linked to lexical decision response times would need to wait for the word to end). It is also compatible with a more simultaneous presentation mode (when one wants to use that instead of our usual sequential presentation mode). It may be revised upon arrival of later information, if need be (e.g. in case of embeddings, late mismatches, garden paths, etc.). Since the IP is only defined for the target word (or target words, in case of a sequence of words in FN5), this means of course that this target needs to be pre-specified to the model, at the beginning of a simulation. Without knowing the correct target in advance, the model will activate exactly the same words and produce the same activation graphs, but it will not compute the IPs. (For nonword inputs, the model will, in general, initially activate the closest candidate words and eventually drop them all, due to the various inhibitory mechanisms.)

Box 6. To get the isolation point of a word or sequence of words.

In BIMOLA, use as target the French verb “boite” (/bwat/) and click “Input all”. Observe that during /b/, /w/, and /a/, “boive” is a higher candidate than “boite” (because the former is more frequent than the latter). However, during the /t/, “boive” is quickly inhibited (by its inactive /v/) while “boite” accumulates more activation (from its /t/) and eventually surpasses “boive”. This moment in the simulation, the isolation point of target word “boite”, is marked with a little flag, labeled “IP”. To get the exact cycle, either click on the word activation graph and hold (use the crosshairs that appear to measure the cycle), or simply click “Best 6” and find the IP@cycle (= 55) and %IP (= 68.75%) written in the log. Note that the IP is not the same thing as the UP (uniqueness point), which occurs, in the case of “boite”, after the end of the word since there also exists “boitille”; but “boitille” is less frequent that “boite” and thus never surpasses “boite” before it is inhibited as well as (during the final silence input).

In FN5, use the sequence “le cygne” (/ləsiɲ/), click “Input all”, finally “Best 6”. The article “le” has a very early isolation point (IP@cycle = 4, %IP = 8.33%) as it is such a frequent word. Observe that “cygne” gets an IP although it actually does not surpass “signe”. These two words are homophones. The definition of the IP demands only that “cygne” surpasses all other candidates (e.g. “signal”, “signature”, “système”, etc.), not also “signe” (of course, it never will). By testing the sequence “le signe” as well, we find that the IP of “cygne” is later (IP@cycle = 82) than the IP of “signe” (IP@cycle = 69): “signe” is the more frequent word of the two and therefore surpasses the other candidates (“signal” and the rest) faster than “cygne” does. FN5 also indicates in its log the “IP@cycle Word” and the “%IP”; both are calculated from the start of the word (i.e. from the /s/).

To get the IP for a whole list of test words (or sequences of words, for FN5), open the “Macros” panel. Click “My.List”. Write the test words (or sequences), line by line, within double quotes, and click “Store”. Within the “Macros” panel, click “Go” to execute the macro named “Batch.MyList.Macro”. The model will automatically run all the test items, one by one. At the end, a window entitled “My.Out” will give the results.

CHAPTER 5. SPECIFIC MECHANISMS

This chapter presents the internal mechanisms that are specific either to FN5 or to BIMOLA. The monolingual model FN5, which recognizes single words as well as sequences of multiple connected words, includes specific mechanisms related to sequential processing: they determine each word's position within a sequence of words, they select a pronunciation variant if the word has several, and they check the words preceding and following the word and its syntactic context. The bilingual model BIMOLA only processes single words but it deals with two languages all at once, therefore it contains specific mechanisms that concern language activation: they account for the bilingual's language modes (i.e. global configurations of the bilingual's two languages) and they permit the recognition of base language words as well as guest words (i.e. code-switches and borrowings from one language into the other).

Sequential processing in FN5

How can a localist connectionist model account for the recognition of multiple words presented sequentially in spoken language? What are the mechanisms that enable it to identify, for instance, the two French words "ta table", in correct order, given the sequence of French phonemes /tatabl/? We have found that the proposals put forward by TRACE and Shortlist (that were made for English, of course, despite the French examples that we will use here) are not devoid of problems, and we have hence developed a new approach.

In TRACE (McClelland & Elman, 1986), all the units (words, phonemes, and features) are reduplicated over time, that is, there is one unit for each successive position in the speech stream, so that a word can be activated at any position. Thus, for example, if the speech stream contains the phonemes /tatabl/, for “ta table”, there is an independent word unit /tabl/, for “table”, at the position of the first /t/, another one at the position of the first /a/, a third one at the position of the second /t/, a fourth one at the position of the second /t/, etc. Actually, these copies of word units are placed even every half phoneme. As it happens, TRACE can replicate some interesting effects that concern multiple-word recognition (e.g. it can parse sequences of phonemes into words, it can establish where one word ends and the next one begins without any cues to word boundaries, etc.) but, as the authors noted themselves, the reduplication approach is not very elegant. Others remarked upon the fact that it makes the architecture heavy and that it does not scale up to a large lexicon (Frauenfelder, 1996; Norris, 1994; Strauss et al., 2007). Moreover, some aspects of it are not psychologically very real; an approach with only one unit for each word would be preferable.

Shortlist (Norris, 1994; Norris et al., 1995, 1997) manages without copies of word units and can recognize multiple words even when operating on a large lexical database. However, to do so, it has to limit the number of words (max. 30 best matching candidates per phoneme position) that are permitted to take part in the competition process (i.e. in the lateral inhibition between words). This approach is questionable because (a) the scoring method used to determine the shortlist of best matching words is too simple (+1 for a matching phoneme, −3 for a mismatching phoneme), (b) the size of the shortlist is predetermined (is this realistic? if it were, which size of shortlist would be correct?) in addition to being restricted (but we know that for inputs such as /ta/, a very large number of words match equally well), (c) the decision whether a word is admitted into or excluded from the list is binary (in contrast to graded word activation values in most other models), and (d) the matching scores are updated only when a new phoneme starts (and not continuously as should be the case). Besides, all this differs sharply and alarmingly from TRACE, which keeps all its words involved, at least potentially, from the very beginning to the very end of the process of word recognition.

To summarize, TRACE freely lets all its words participate in the process at all times but, to do so, needs to reduplicate words over all potential positions, including unlikely ones. Shortlist, by contrast, does without multiple copies of words but it has to restrict the number of words that are permitted to take part in the process at any moment to just a few. Since neither of the two solutions is ideal, we have introduced a new approach (and compromise), in which as many words can be active simultaneously as in TRACE, but which functions without reduplication like Shortlist. First described in Léwy et al. (2005), our approach contains a position processor, that is, a processor specifically in charge with the complex task of positioning words in time, which, as we will see, continually computes and chooses the best alignment position for each word.[24] In addition, in case of words with pronunciation variants (i.e. French words with schwas, adjective forms, etc.), our approach uses groups of connections in order to select the best variant. We will finally see how the position processor can take account of some simple (but powerful) constraints that concern the context.

The position processor (PP). At each cycle, PP aligns each word with each phoneme position in the string of phoneme units (position 1 corresponds to phoneme 1, position 2 corresponds to phoneme 2, etc.) and it reads off the word's net input value at each position: the sum resulting from the mechanisms of word activation and inhibition (bottom-up, lateral, as well as from preceding/ following word context; for the last mechanism, see below). PP compares the net input values with one another and chooses the word alignment that gives the highest value. This value is conveyed to the word unit, which converts it to a new activation level (using an activation function that also takes into account the unit's previous activation level). Both the activation level and the alignment position (a temporal position index) are stored in the word unit. This is repeated

[24] We herein adhere to the traditional view of phonemes and words as rather simple counting units that are passive receivers and senders of activation and inhibition (see e.g. Morton, 1969). An alternative view would be to grant word units some form of self-activity and intelligence, as brought up as early as in Selfridge's (1959) Pandemonium model and, more recently, in simulations with multiple autonomous agents (e.g. de Boer, 2000; Cangelosi, 2005). Therefore, instead of assuming that the choice of best position for each word is determined by a special processor, we could postulate that the words have the responsibility to actively search their best position on their own. Since our position processor (or PP) operates separately on each word (i.e. never across several words), the outcome would be the same.

at each cycle (16 per phoneme, in the normal speech rate, as we have seen) and for each incoming phoneme. Naturally, in a sequential (beginning-to-end) presentation mode, there is only a position 1 alignment, and hence merely one net input value, during the first phoneme (normally cycles 1–16). In the course of the second phoneme (cycles 17–32), position 1 and position 2 values are compared; during the third phoneme (cycles 33–48), position 1, 2, and 3 values are compared, and so forth.[25]

For example, for the French phoneme sequence /tatabl/ ("ta table"), the word /tabl/ has a given net input value and a position index of 1 during the first phoneme (i.e. when the /t/ is activating words with an initial /t/, such as /tabl/); it is, de facto, the preferred value since there are no other values to compare it to. During the second phoneme, PP compares the net input value of the word /tabl/ when it is aligned in position 1 (i.e. with /t/) and in position 2 (i.e. with /a/) and finds that the value is higher in position 1 than in position 2 (because /tabl/ does not contain an initial /a/); it therefore stores the position 1 value (in the form of an activation level; see above) and the position index of 1. During the third phoneme (the second /t/), PP compares the net input values for position 1, position 2, and position 3 and finds, after a few cycles within the /t/, that the position 1 value is now lower than the position 3 value (due to the mismatching second /t/ and the inhibition from absent /b/ of /tabl/); PP therefore stores the position 3 value as well as the corresponding position index of 3. What has just happened is that the word /tabl/ has shifted from being a candidate aligned with the beginning of the phoneme sequence (/ta····/) to being a candidate aligned with the position 3 (/··tabl/). This is just what one would want since /tabl/ cannot continue to stand as a candidate for /tat/ but it can be a new (perfect) candidate for the sequence beginning with the second /t/. From then on, the position index for the word /tabl/ will remain the same (i.e. position 3) since a comparison of net input values in other positions shows that this is indeed the best alignment possible for it. Of course, the first word of the sequence, /ta/, has increased its activation level meanwhile (it made additional progress when /tabl/ moved over to position 3 and stopped competing with it). In the end, both "ta" (in position 1) and "table" (in position 3) come out as the top candidates for the string /tatabl/. A sequence of two words has been recognized!

---

[25] In the fast speech rate (with 3 cycles per phoneme), things happen more quickly (one position value during cycles 1–3, two position values during cycles 4–6, etc.). In an entirely simultaneous mode (in the manner of visual word recognition models), all position values would be compared from the process's very beginning (cycle 1).

Box 7. To study the position processor (PP) in FN5.

Take the example of “ta table”. But first, use the “Parameters” panel to turn off “Context: prenominal word at beginning”, so as to permit that nouns like “table” can occur in all positions. Type the target words “ta table” into the “Words” field. Enter the sequence of phonemes, /tatabl−/, phoneme by phoneme, by clicking “Input 1”. Also click “Best 6” after each phoneme. As we explained just above, during the first /t/ and first /a/, “table” is a candidate aligned with position 1 (i.e. as written in the log, “table” extends with its four phonemes over positions 1–4). From the second /t/ forward, “table” shifts to position 3 (i.e. it now extends over positions 3–6). Indeed, this is where “table” belongs, with the result that it nicely goes with “ta”, which has held on, all the time, to position 1 (i.e. it extends with its two phonemes over positions 1–2).

In the top part of the word activation graph, the phonemes already entered are indicated along with their positions to clarify the correspondence. To label also the words with their position index, open the “View” panel and click “s[ ]”.

Evidently, PP implements a process of constant optimization (which is, for this word, the current best alignment position that yields the highest net input value?). The process operates continually, at each cycle, and hence the word’s best alignment position can change from one cycle to another. After each cycle (i.e. once the new highest net input value has been determined), one activation level and one position index are stored in each word unit. Contrary to Shortlist, all the words in FN5 are involved in the process simultaneously. Though many words will have, of course, an activation level below 0 and thus remain inactive (i.e. without influence), the number of the words that reach an activation level greater than 0 is completely unrestricted in FN5 and it can change dynamically. For example, as long as FN5’s input is only the initial /t/ of /tatabl/, there are still hundreds of possible outcomes (in FN5’s lexicon, 907 words begin with /t/, and the words to be recognized by the model could be just any of these 907 words). But when FN5’s input becomes the complete sequence /tatabl/, just two words remain (“ta” and “table”). If it is appropriate to the word recognition situation, a multitude of words can be activated in FN5 in parallel, without any problems (as it is typical for localist connectionist models). There is no need at all to artificially constrain the number of candidates in advance, as it is done by Shortlist[26] (this rather reminds us of the tradition of serial processing; see e.g. Forster, 1976).

[26] Shortlist B (Norris & McQueen, 2008) holds on to the same attitude.

Every word has only one position, at a time, in FN5. In that sense, there are no reduplications of word units as in TRACE (which retains all the time the word activation levels at all the positions, including many unlikely ones). In most cases actually, a word's position is clear and unambiguous, and there is no use for keeping the word alive in several alternative positions, in which it does not occur right now (as it is done in TRACE). In the example of /tatabl/ ("ta table"), the word /tabl/ switches between two positions, but there are no further ones. Take the example of /taʃɛz/ ("ta chaise"), in which the word /ʃɛz/ is sure to go into position 3 (i.e. beginning with the /ʃ/) and nowhere else (thus, during the input of /ta/, the activation level of /ʃɛz/ is close to the minimum activation level). Spoken words can, of course, appear in any temporal position imaginable; nonetheless, it suffices to represent each word by just one word unit and then have PP find out the position where, according to the net input value at that position, the word occurs this time (or, at least, where it is most likely to occur, if anywhere). Should several positions produce exactly the same net input value, PP favors the most recent position over earlier ones. This is important for word repetitions such as /tabltabl/ ("table table"): initially, "table" at position 1 will be proposed, but sometime later, "table" will be transferred to the position 5 (i.e. position 1 shifted by four phonemes). We should note that as isolated words and in the rather short sequences that we currently use (determiner + noun, prenominal adjective + noun), words are not repeated; therefore, this aspect has not been very relevant so far.[27]

<u>*Groups of connections for variant selection.*</u> Words with phonemes that can be deleted (e.g. a schwa vowel) and/or added (e.g. a liaison consonant) are assumed to have separate groups of phoneme to word connections for each of their pronunciation variants (for a discussion of FN5's words with pronunciation variants, which also include contracted words like the article "l' ", see pp. 43ff). For example, for the noun "fenêtre", there is one group of connections for the

---

[27] As a possible future extension of PP to generally treat words that reoccur in the same sentence or in later sentences, we suggest to keep in each word a log of the word's past recognition positions (e.g. "table" would keep track that it was isolated by PP at positions 1, 5, 17, 89, etc.). These log entries or past recognition positions of a word would be purely historical and inactive (unlike TRACE's multiple active positions of the same word). How many previous positions a word had could inform the model on the word's frequency of occurrence and would theoretically make our current stored word frequency values dispensable.

pronunciation variant with the schwa absent (i.e. /fnɛtʁ/) and there is one group representing the variant with the schwa present (i.e. /fənɛtʁ/). Likewise, for the adjective "grand", there is a group of connections for the feminine form (/gʁɑ̃d/), one for the masculine form without liaison (/gʁɑ̃/), and one for the masculine form with liaison (/gʁɑ̃t/). PP aligns each word with each phoneme position in the phoneme string (position 1, position 2, etc.) but, in the case of several groups of connections, it reads off the net input value at each position for each group of connections so as to determine the combination of position and group of connections (i.e. pronunciation variant) that has the highest net input value (this might currently be e.g. the variant with schwa absent, when at position 4). At the end of the comparison, information regarding which group of connections and which position produced the highest net input value is stored into the word. This process is repeated at each cycle; after each cycle, a word is represented by only one group of connections and one position.

The linguistic knowledge of which groups of connections are available for a word is memorized explicitly in the word units since the number and the kind of variants differ for each word (one word may have two optional schwas and a liaison consonant represented by eight groups of connections, but another word may have just one group of connections). The concrete connections, however, physically joining the phonemes to the words, are not hard-wired in advance, but rather soft-wired dynamically, that is, these links are entirely built on the fly, depending on the variant selected and the position concerned. This permits to keep down the quantity of connections in FN5 to just those presently needed. McClelland's (1985) connection information distributor (CID) was a somewhat comparable mechanism; it allowed to represent general connection information in one hard-wired, central store and to load it into soft programmable modules, depending on the situation; it was used in McClelland's (1986) programmable blackboard model (PABLO) of visual word recognition.

We note that FN5 binds a number of indices or variables, time and again (i.e. once in every cycle), to their current optimal value. All the words are being bound to their best alignment position within the phoneme input string. Words with pronunciation variants (i.e. schwas being deleted or not, liaison consonants being present or absent, adjectives ending in feminine or masculine forms, etc.) are also being tied to their best group of connections (which actually is nothing but an index, stored in the word units, like the index of position). By this process of binding, word recognition is made invariant of the actual position and group of connections. Binding is an elementary function that has been hypothesized even in research and theory on real neurons in the brain (e.g. Engel & Singer, 2001; von der Malsburg, 1999); this should mean an argument to allow for this concept in a localist connectionist model of human behavior.

Attenuating the lateral inhibition for ending variants. In the construction of FN5, we ran into an interesting puzzle regarding word-level lateral inhibition in French. It concerns the consonant to the left of the boundary between words (e.g. the /k/ in the phoneme sequence /ʃakuʁs/ of "chaque ours"), which was able sometimes to incorrectly activate two words that shared this consonant (e.g. "chaque course", /ʃak/ and /kuʁs/), when the word-level inhibition (between the partly overlapping "chaque" and "course") was not high enough in FN5. Then again we desired that, in the case of a consonant of liaison (e.g. the floating /t/ in /pətitami/ of "petit~ ami", a phoneme sequence that can also be segmented into /pəti/ and /tami/, "petit tamis", i.e. without liaison but with the consonant being attached to the right of the word boundary), FN5 should let activate not only "ami" and "petit" (ending with /t/) but also, at least momentarily, "tamis" and other words starting with /t/, like "tapis", "tas", and "tableau". However, FN5 showed what we intended (a temporary ambiguity of where to place the word boundary) if and only if word-level inhibition was kept relatively low. The situation, manifestly, was that one requirement clashed with the other!

To solve this puzzle, we have, on the one hand, increased the general parameter of lateral inhibition between words (it is −0.15 per phoneme position). In this way, the input /ʃakuʁs/ now correctly activates "chaque ours", no longer "chaque course", because the words "chaque" and "course" now inhibit one another mutually, and hence prevent a double use of the /k/. On the other hand, we have attenuated the lateral inhibition of words for the phoneme positions of ending variants, and have introduced a divisor (a new parameter with a value of 6.5), in the sense that words like "petit(e)", with a liaison /t/ at their end, now inhibit words like "tamis" or "tapis", that contain that /t/ at their very beginning, this much less. The attenuation only concerns the word's ending phoneme position (the one containing the liaison consonant) and it is strictly one-way (i.e. "tamis" and "tapis" inhibit "petit(e)" just normally, without any attenuation). On the connectionist level, these phoneme strings are separated into a main part (/p(ə)ti/) and extension (/t/), with different outgoing weights of lateral inhibition toward other words (but with the same incoming weight). As a consequence, we have enabled FN5 to show the typical difference, established experimentally (Yersin-Besson & Grosjean, 1996), between linking without and with liaison: Phrases like "chaque ours" (linking without liaison) are recognized quickly and without competition by words starting with /k/ like "course", while phrases like "petit~ ami" (linking with liaison) are usually recognized as well, but only so after a passing moment of lexical ambiguity (since various words beginning with /t/ are also temporarily activated), and therefore often with a certain delay.

Furthermore, FN5 predicts a performance quite similar to liaison for some other pronunciation variants that we have discussed: those concerning contracted forms (e.g. "c'te" vs. "c' ", "notre" vs. "not' "), feminine forms (e.g. "grande" vs. "grand"), and plural forms (would the lexicon contain them, which at this time it does not; e.g. "grandes"). According to FN5, also these ending variants, like those for liaison, are to be expected to often generate a transient state of ambiguity, and hence hesitation, of where to place the word boundary (which now and then may lead to a small recognition delay). In fact, in FN5, the lateral inhibition is attenuated (i.e. divided by 6.5, as for the liaison consonants) for the ending phonemes of all these variants (e.g. for the /t/ in /stafɛʁ/ "c'te affaire", the /ʁ/ in /nɔtʁadjø/ "notre adieu", the /d/ in /gʁɑ̃dane/ "grande année", and the /dz/ in /gʁɑ̃dzikon/ "grandes icônes"). This, of course, allows words that do not end, but start, with the underlined phonemes to become activated too, which is a good thing since sometimes they may be the ones to be recognized (e.g. when the input is /stapi/ "c'tapis", /nɔtʁadjo/ "not' radio", /gʁɑ̃danwa/ "grand danois", or /gʁɑ̃dzigan/ "grands tziganes"). It would be interesting to study, and experimentally test, FN5's prediction of a momentary hesitation for these kinds of sequences (e.g. in a gating study, using stimuli that by no means give away any acoustic clues about the word boundary).

Influence of preceding and following context. We have seen so far the general word activation and inhibition mechanisms that occur during a word: The word gets activation from the phonemes it contains that are present, and it receives inhibition from the phonemes it contains that are absent. However, what comes before the word, and what goes after it, are very important as well! By and large, a spoken word is rarely identified just by itself, that is, one word at a time (cf. the now outdated views of Cole & Jakimik, 1980; Marslen-Wilson & Welsh, 1978). On the contrary, a word is usually recognized in (and together with) a surrounding context of adjacent, connected words, both prior to the word and subsequent to it (the role of later information being known at least since Grosjean, 1985). This preceding/following word context may facilitate as well as impede the recognition process greatly. Therefore, FN5 accounts for it by two specific mechanisms: an extra bottom-up activation (a reward, so to speak) when at least one word is activated directly preceding the word, and an extra bottom-up inhibition (a penalty) whenever it happens that no words at all are activated that immediately follow (i.e. continue) the word.

The preceded words activation mechanism is implemented as a small activation bonus from first phoneme to word (parameter value: 0.35), in case that the word is preceded by one word or several words that end exactly one phoneme position earlier. The mechanism is helpful for phrases with elisions and contractions, like “l’objet” (/lɔbʒɛ/), “l’ouvrage” (/luvʁaʒ/), and “c’portant” (/spɔʁtɑ̃/), which are recognized correctly as long as the mechanism is turned on, but which would be parsed into *“lobe jet”, *“louve rage”, and *“sport temps” (phonemically the same but with another word boundary as intended), when this mechanism is turned off. As a matter of fact, the mechanism is an effective means to counterbalance FN5’s inclination to prefer, partially at least, longer words (i.e. “lobe”, “louve”, “sport”) over shorter words (the “l’ ” and “c’ ”); a normal and natural tendency, which other models like TRACE and ARTWORD show as well, and which makes FN5 sometimes try to continue attaching later arriving phonemes (e.g. the /ɔ/ and /b/ of “objet”) to the end of the first word (“lobe”). With the preceded words activation turned on, FN5 rather gives some head start to the second word (e.g. “objet”, if preceded by “l’ ”), which may just help that word win the competition (over “lobe”, and together with “l’ ”).

Box 8. To examine the role of “Preceded words activation” in FN5.

To see the aforementioned behavior clearly, keep “Context: number of words” and “Context: prenominal word at beginning” turned off, and switch “Preceded words activation” alternatively on or off. Then use “l’objet”, “l’ouvrage”, or “c’portant”.

Another case in point are phrases that contain a word with an aspirated *h*, like “quelque houx”. These are normally produced with a disjunction (Grevisse, 1993; Tseng, 2003), that is, a short hiatus or glottal stop, which is transcribed in FN5 as a silence segment (i.e. /−/), between the words (/kɛlk−u/). Owing to the phoneme position taken up by the silence, words cannot directly precede “houx”. Instead, the preceding silence is taken into account for the recognition of these words; its presence gives an additional activation, and its absence (as in /kɛlku/ of “quel cou(p)” or “quel coût”) results in an inhibition of “houx”. This aspect of the preceded words activation mechanism is also relevant for sequences like “ma jupe”, “ta salle”, “la r’vanche”, and “la s’maine” (the last two with deleted schwa), which should in no way activate, respectively, *“mage huppe”, *“tasse halle”, *“larve hanche”, and *“l’asthme haine”. (Those phrases do not present a problem to FN5 as long as word frequency is in effect.)

The words without continuation inhibition[28] is a binary (not a numerical) setting; normally, this mechanism is on, but it can be turned off if one wishes to examine how the mechanism influences the recognition process (see Box 9). It pertains to short (i.e. mono- and disyllabic) candidate words that are embedded in longer words to be recognized (such as “voix” in the beginning of “voiture” or “voisine”, and “quart” in the initial part of “carriole” or “carpette”) and, even more crucially, those astride two words (e.g. “macabre” in “ma cabriole”, “vieillard” in “vieil arbre”, “montre” in “mon truc”, “monde” in “mon duc”, “mongol” in “mon golfeur”, “lapin” in “la peinture”, “lac” or “laque” in “la queue”, “lapereau” with its schwa deleted to “lap’reau” in “la prose”, etc.). If we turn the mechanism off, these erroneous candidates tend to keep their early strong activation without being affected by the arrival of the rest of the phoneme sequence. But when we turn the mechanism on, these erroneous candidates are bottom-up inhibited (and hence effectively excluded from being candidates) as soon as FN5 detects that they do not have a so-called continuation, that is, when these words cannot be continued to a sequence of words in any possible way. For instance, in the case of “voix” (/vwa/) in “voiture” (/vwatyʁ/), FN5 will find that no continuation is activated beginning in the position of the /t/ that follows “voix” (indeed, a word /tyʁ/ does not exist in French, therefore a sequence that starts with “voix” is not possible), and hence it will deactivate “voix”. For “macabre” (/makabʁ/) in “ma cabriole” (/makabʁijɔl/), FN5 will discover that no word is activated starting in the /i/ position (as the lexicon does not contain any word beginning with /ijɔ…/), so it will inhibit “macabre” (likewise for all the other examples). The deactivation of a wrong candidate word does not take place immediately at the word’s offset (at that moment, FN5 cannot know yet what will possibly follow), but it occurs usually a short time afterward (two phonemes later at the earliest, in the current implementation, i.e. during the /j/ of /makabʁijɔl/, and during the /y/ of /vwatyʁ/).

[28] In BIMOLA, which operates only on isolated words (verbs), the silence that follows a word has a pivotal role for short words that are onset-embedded in longer words (e.g. “fasse” in “façonne”, “note” in “notifie”, “parle” in “parlemente”, and “passe” in “passionne”, “pacifie”, or “pasteurise”). Assume that “façonne” (/fasɔn/) is presented to BIMOLA. For the candidate “fasse” (/fas/), BIMOLA expects a silence after the /s/. When there is none (since the /ɔ/ of “façonne” is entered instead), the absence of the silence creates a bottom-up inhibition to “fasse”; hence “fasse” is deactivated, while “façonne”, of course, continues to get bottom-up activation and is recognized by BIMOLA, as it should be. FN5’s words without continuation inhibition is a more generalized version and extension of this mechanism to multiple-word sequences.

For some words, it may come about quite tardily, that is, several phonemes into the next word or words. We note that even the longest of words, which typically contain numerous embedded short words, can be recognized in this manner. For example, “universalisation” boasts as many as 27 embeddings in FN5’s standard French lexicon (24 in its Swiss French lexicon), including “une”, “uni”, “univers”, “verre”, “sa”, “salle”, “lit”, “as”, “ion”, etc., but no single combination of embeddings can make up the complete sequence /ynivɛʁsalizasjɔ̃/, and so they all will be inhibited at some moment.

More generally, by inhibiting words without a continuation, FN5 can find its way out of the dead ends, or garden paths, of which spoken word recognition is strewn with over and over again (Grosjean, 1980), particularly so, as it is the case for FN5, on a level of purely phonemic sequences and without recourse to any explicit acoustic cues to word boundaries (cf. Davis, Marslen-Wilson, & Gaskell, 2002). We have found it fascinating to closely observe simulations with these types of words. In all the examples mentioned, FN5 steers momentarily into a garden path (i.e. activates a wrong word candidate during some time), but it can get out of it (i.e. inhibits that candidate later on), and it ultimately comes up with the right result, after all. Interestingly, in the case of garden paths with two-word sequences (e.g. “vieil arbre”, “mon truc”, etc.), the two correct short words (i.e. “vieil” and “arbre”, “mon” and “truc”) are very often isolated by FN5 at one and the same moment: As soon as FN5 has deactivated the garden-path word, both the first and second of the correct words have their chance to get sufficient activation and are then usually recognized in the same (or a nearby) processing cycle. It means that the first word is perceived very late, that is, well after its offset (for experimental evidence on the short words’ late recognition, see Grosjean, 1985; Bard, Shillcock, & Altmann, 1988).

Box 9. To examine the role of “Words without continuation inhibition” in FN5.

To replicate the behavior described above, turn off “Context: number of words”, “Context: prenominal word at beginning”, and “Word frequency” (i.e. remove the check marks next to each of the three parameters and click “Apply”).

Use some of the examples mentioned, say, “vieil arbre” (/vjɛjaʁbʁ−/). Observe that “vieillard” (/vjɛjaʁ/) becomes an excellent candidate during /aʁbʁ/ of “arbre” but that it is promptly inhibited during the silence that follows. FN5 discovered, thanks to the “Words without continuation inhibition” mechanism, that “vieillard” cannot stay a candidate after all, because the phonemes that follow (i.e. /bʁ−/) do not activate any word. At this time, “vieil” and “arbre”, the two target words, can become activated and are isolated soon (IP@cycle = 139, for either word). That is, FN5 went into a dead end (by proposing “vieillard”) but it successfully made its way out of it (by ultimately activating and isolating “vieil” and “arbre”). Now also switch off “Words without continuation inhibition”, and test “vieil arbre” again: “vieillard” remains the best candidate until the end of the simulation! Apparently, this mechanism is indeed needed.

Note that when we turn on “Context: number of words”, “Context: prenominal word at beginning”, and “Word frequency”, there are no garden paths anymore (because “vieillard” is not two words), and very frequent words (like “vieil”), even when short and embedded, are activated early and quickly, without a problem. This is the default setting of FN5.

We should remark that our mechanism of words without continuation inhibition is based on a dynamic, lexical, and hence language-specific criterion (do in this language’s lexicon any words exist that the input currently activates in such a way that they follow and continue the word in question?). It therefore contrasts very much with the approach by Norris et al. (1997), who proposed (and implemented for English in Shortlist) an entirely static, phonotactic criterion, called the PWC, or possible-word constraint (does the unused rest of the input between the word in question and the nearest likely word boundaries, before and after the word, contain any vowel or only consonants?). The PWC was claimed to be language-universal (Norris, McQueen, Cutler, Butterfield, & Kearns, 2001) but, at least for French, it runs into a first problem when it is applied to vowel-initial words following liaison or linking without liaison, since this would leave a consonant alone (e.g. the liaison /ʁ/ of “dernier~ oignon”, and the non-liaison /k/ of “chaque avion”) and, unless one assumes to switch the PWC off for just these cases, would interdict the recognition of the vowel-initial word (“oignon” and “avion”; cf. Gaskell, Spinelli, & Meunier, 2002; Spinelli, McQueen, & Cutler, 2003). There is a second problem for the PWC: As we have seen when discussing words that can be contracted, French contains three spoken word forms, all of them frequent determiners, that are composed of just a single consonant (the /l/ in “l’été”, the /s/ in “c’truc” and “c’monsieur”) or two consonants (the /st/ in “c’t acteur” and “c’te actrice”). Here again, the PWC would penalize and hence prevent the recognition of the following nouns (“été”, “truc”, etc.), claiming the isolated consonants were not a possible word. While our proposal (i.e. the words without continuation inhibition) bears some passing resemblance to the PWC (in the sense that both these mechanisms reduce the activation of certain incompatible words) and while it may occasionally penalize candidates just when they leave single consonants stranded (e.g. “bout” /bu/ when “bouc” /buk/ is presented, or “chaud” /ʃo/ in the case of “chose” /ʃoz/, etc.), our approach differs from the PWC in that it concerns the presence (and actual activation) of words in the particular lexicon. For French, as a consequence, it does certainly not prevent the recognition of linked words (e.g. “dernier~ oignon” and “chaque avion”, the liaison case occasionally being somewhat delayed in FN5 by reason of the attenuated lateral inhibition, as we have seen), nor does it have any problems with elisions and contractions (e.g. “l’été”, “c’t acteur”, etc.).

Influence of syntactic context. Once we start to model the recognition of a sequence of words, as we do indeed in FN5, we must accept that we enter the domain of higher level processing where the role of the context—not just of the preceding and following words, but the syntactic or even semantic context—would need to be considered (see e.g. Magnuson, Tanenhaus, & Aslin, 2008; Marslen-Wilson, 1987; Marslen-Wilson & Tyler, 1980; Mattys, Melhorn, & White, 2007; Moss & Marslen-Wilson, 1993; Salasoo & Pisoni, 1985; Tyler, 1984).

Although we do not introduce these levels in FN5’s architecture by any explicit means or representations, the position processor (PP) is already a first step in that direction. In fact, PP can be made to check whether the word would be correct, in an alignment position under consideration, regarding the number of words that are to be recognized (i.e. a single word, a sequence of two words, etc.) and/or regarding the lexical category (the first word of two-word sequences can be required to be a determiner or prenominal adjective[29]; e.g. “le foulard”, “beau foulard”). PP then dismisses a particular position for a word if it is unable to fit the given constraints (e.g. an additional word in the case of single words, or an initial noun when there should be a prenominal word at the beginning, etc.). Of course, PP can also be operated with the word number check switched off (then, the number of words is totally free) or with the lexical category check switched off (in that case, the first word of a two-word sequence is allowed to come from any lexical category). Two binary parameters determine the mode of operation; they are called “Context: number of words” and “Context: prenominal word at beginning”. By default, they are both turned on, but we will also evaluate FN5 when they are turned off.

[29] To permit sequences longer than two words (e.g. “un beau, chaud, long foulard”), the constraints are formulated as follows. Initial word: determiner or prenominal adjective; middle word(s), if any: prenominal adjective(s); final (or single) word: prenominal adjective or noun. (These longer sequences will not be evaluated.)

We argue that the presence or absence of such contextual constraints reflect, either a vital part of the spoken word recognition process as such, or else, at least, the various modes in which listeners in a spoken word recognition experiment find themselves, depending on the stimulus materials used. In fact, participants typically get to know what kind of words to expect if they are being tested on single words all the time (or on multiple-word sequences of always the same build-up), but they do not when they are being tested on syntactically mixed stimuli. Therefore, FN5 too should reflect whichever contextual mode it is supposed to account for. Certainly, by adding these two simple (but powerful) constraints, we can give a further boost to FN5's recognition success rate. But, as we will demonstrate later in the FN5 evaluations, that rate is very high even without the contextual constraints.

Language activation in BIMOLA

Experimental research on bilingual spoken word recognition has accumulated a large body of results. Apart from the authoritative work of Grosjean (1988, 1997, 1998, 2008; Soares & Grosjean, 1984), additional studies have been conducted more recently (e.g. Blumenfeld & Marian, 2007, 2013; Chen, 2008; Dornbusch, 2012; Hernandez, Bates, & Avila, 1996; Ju & Luce, 2004; Lagrou, Hartsuiker, & Duyck, 2011, 2013; Li, 1996; Li & Yip, 1998; Marian & Spivey, 2003; Schulpen, Dijkstra, Schriefers, & Hasper, 2003) and thus certainly call for a computational model of this process. This is precisely the role of BIMOLA, our English–French bilingual model.

Accounting for the bilingual's language modes. Bilinguals go in and out of several language modes, that is, global configurations of their two languages (Soares & Grosjean, 1984; Grosjean, 1997, 1998; Léwy & Grosjean, 2008). Which one it is, at the present time, depends on the conversational situation. They are in a monolingual language mode when hearing monolinguals who speak only one of their languages (one language is used in the conversation); they are in a bilingual language mode when they listen to other bilinguals who speak the same two languages and who feel at ease mixing them by code-switching or borrowing words from one into the other language (both languages are used the conversation). At any moment in time, one of the bilingual's two languages serves as the current main language of the conversation; it is called the base language (or matrix language, Myers-Scotton, 2002). As we depict in the top part of Figure 11 (which we drew by adapting and extending the one by Grosjean, 1997), its choice is discrete: either language A or language B, that is, either English or French, is the base language. While the base language is always at 100% activation, the other language, called the guest language (or embedded language), can find itself at various levels of activation, from 0% to 100%. This is visualized, in the middle part of the figure, by a continuous line (language mode continuum). The position of the dot on this line denotes what the language mode currently is (rather monolingual, or rather bilingual). Only very rarely, if ever, do bilinguals have both their languages activated at 100% (because, in any conversation, one of the languages prevails); therefore the continuum ends, on its right, in dashes. Finally, in the bottom part of our figure, the two languages' levels of activation are shown graphically by the height of two bars, one for the base language and the other for the guest language (the bars break off, on their right, at an assumed maximum of about 80% for the guest language, since a higher level of activation than this is unlikely). The

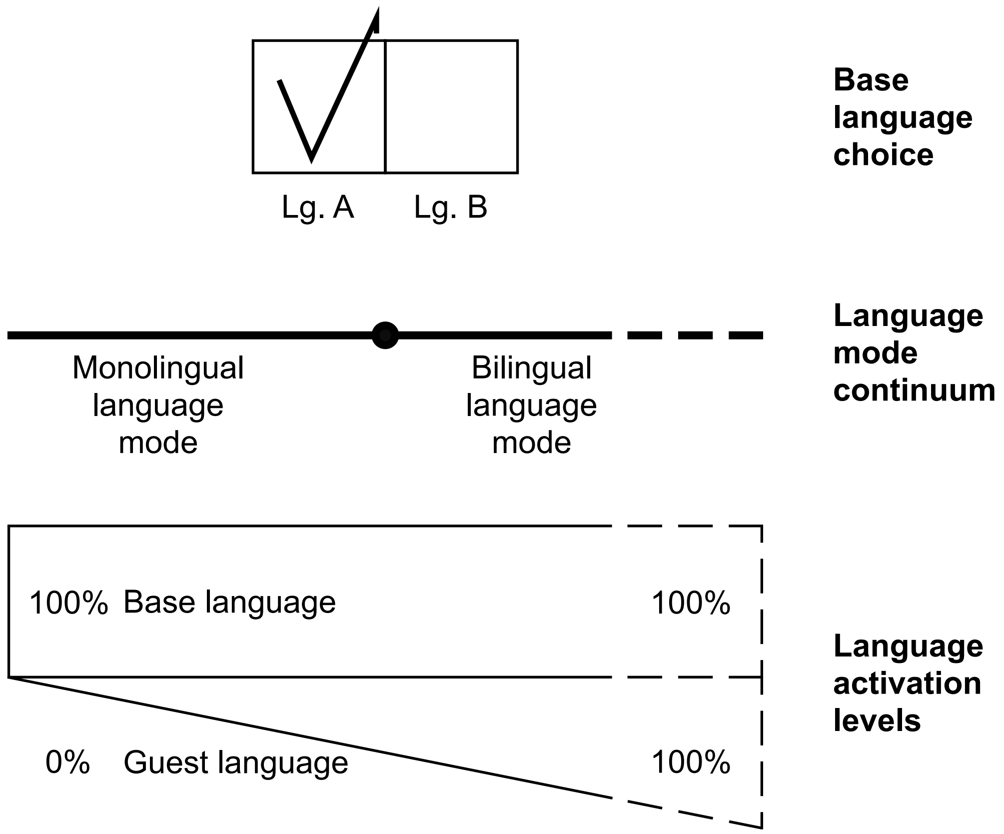


Figure 11. Schematic representation of the bilingual's language modes: base language choice, language mode continuum, and language activation levels.

language mode is set once in advance (according to the present conversational situation) but, of course, may change from time to time (in the event of a new conversational situation).

BIMOLA can operate in any of these language modes and it does so by means of higher or lower resting levels for all its units (phonemes and words). As we already mentioned, the resting level of a unit is the initial activation level, from which a unit in the localist connectionist model commences and to which it tends to decay back (McClelland & Rumelhart, 1988). The units of the base language all start from a resting level of 0 (i.e. the neutral resting level, which is used also for all units in our monolingual model FN5). The units of the guest language, however, begin from a lower resting level, somewhere between 0 down to −1 (the usual minimum activation level of localist connectionist units). They are slightly negative but still close to 0, should the bilingual listeners find themselves on the right-hand side of the language mode continuum (i.e. in a bilingual language mode); but the guest language units are even lower (more negative), if the bilingual listeners are on the left-hand side of the continuum (i.e. in a monolingual language mode).[30] In any case, compared to base language units, guest language units will, as a result, always need more activation (and/or less inhibition) to become active. Guest language phonemes will take more time before they can activate words, and guest language words, as well, will need more time until they are finally isolated. The units' exact resting levels depend, of course, on the position (the dot) on the language mode continuum and the respective language activation level L (always 100% for the base language, and 0% to about 80% for the guest language). We employ the following nonlinear function with an exponent γ. The higher this parameter, the larger the effect of language mode. It is high for words (γ = 0.9) but less so for phonemes (γ = 0.1); thus, guest language phonemes will usually still have quite a good chance to be activated (even in the monolingual language mode) but guest language words will have a harder life at times (particularly in the monolingual language mode):

$$\text{rest}(L) = -1 + \left(\frac{L}{100}\right)^{\gamma} \qquad (8)$$

---

30 Setting the resting levels of all the phonemes and words of a language, once in a while, is entirely different from having, as in BIA (Dijkstra & van Heuven, 1998; van Heuven et al., 1998), a language node connected to all the words of that language, continually summing up the activity of the words in the one lexicon, and repressing, by means of inhibitory links, the words of the other lexicon.

We will have BIMOLA run with both either English or French as the base language but, even though we are at total liberty to choose the activation level of the guest language, we suggest to normally work with the levels of 80% (for a bilingual language mode), 50% (for an intermediate language mode), and 20% (for a monolingual language mode). At a guest language activation level of 0%, BIMOLA runs, de facto, with only one language (one set of phonemes and one set of words) and becomes a purely monolingual model, quite similar to our other model, FN5. The other language (set of phonemes and set of words) is physically still there, but it is dormant and is not accessed.[31]

Box 10. To set the language mode in BIMOLA.

Open the “Language mode” panel (via its top-left box). Use the radio buttons to select either English or French as the base language. Move the horizontal slider to set the activation of the guest language to any level between 0% and 100% (but keep in mind that 100% guest language activation is very uncommon).

Eight presets are provided: “E biling.” (bilingual mode, English as the base language, French as the guest language at 80%); “E int’med.” (intermediate mode, English as base language, French at 50%); “E monoling.” (monolingual mode, English as base language, French at 20%); “E only” (English only, French completely deactivated); and, vice versa, with French as base language and English as guest language: “F biling.”, “F int’med.”, “F monoling.”, “F only”. (N.B. When BIMOLA is newly launched, it is in “F biling.” language mode, i.e. has French as base language and English as guest language at 80%.)

As an example, choose French as the base language and use the English verb “snoop” as target. Set the guest language activation to 80%, 50%, and 20%, and, for each case, click “Input all”. Note that French candidates show up first (since French is the base language and therefore is always activated at 100% language activation). As for the English candidates, including the word “snoop”, they show up earlier in case of 80% guest language activation (bilingual mode) than for 50% guest language activation (intermediate mode). For the case of 20% guest language activation (monolingual mode), one has to have BIMOLA run additional simulation cycles after the end of “snoop” (by clicking “Cycle”) to make that word appear at long last. The IP@cycle of “snoop” is, respectively, at cycle 52, 59, and 93. Lastly, set the guest language activation to 0%: English is

31 See the studies on language attrition in bilinguals (e.g. de Bot & Stoessel, 2000; Schmid, 2010).

now totally deactivated and “snoop” will never be activated (even with hundreds of times clicking “Cycle”). Thus, BIMOLA becomes a purely French monolingual model, comparable to our other model, FN5.

Try it the other way around, with English as the base language and French at 80%, 50%, 20%, and 0% guest language activation. Of course, the target word “snoop” is now a base language word and is activated identically in each of the four cases (with IP@cycle = 48).

To visualize words below 0 activation level, open the “View” panel: uncheck the check box next to “only part > rest. level” and change “Y min” to −1. Turn off the parameter “Word frequency” to purely see the role of guest language activation.

Recognizing guest words, with or without guest language pronunciation. Bilingual listeners are able to identify auditory words from either of their two languages. They recognize, of course, words that are part of the base language, but they can also recognize words that come from the guest language. These guest words can be brought in dynamically into the base language in two ways, either by code-switching or by borrowing (Grosjean, 1988; Myers-Scotton, 1992, 2002; Pfaff, 1979; Poplack, 1980; Poplack, Sankoff, & Miller, 1988). Take the following English/French mixed-language utterances (from Grosjean, 1988):

(1)	Il faudrait qu’on	drop	nos deux cours.	CS: E /dɹɒp/	B: F /dʁɔp/
	“We should	drop	our two classes.”		
(2)	Il faudrait qu’on	tag	ces cinq jupes.	CS: E /tʰæg/	B: F /tag/
	“We should	tag	these five skirts.”		
(3)	Il faudrait qu’on	peel	toutes les pêches.	CS: E /pʰiɫ/	B: F /pil/
	“We should	peel	all the peaches.”		

In each example, an English word (a verb in the base form) is brought in as a guest word into a French sentence. If that guest word (“drop”, “tag”, or “peel”) is pronounced using guest language phonemes (i.e. English /dɹɒp/, /tʰæg/, and /pʰiɫ/), it is a code-switch. The same guest word can also be pronounced using base language phonemes, that is, it is adapted phonetically into French (e.g. as French /dʁɔp/, /tag/, and /pil/, or comparable); in this case, it is a borrowing. (To separate from lexicalized loans, the latter is sometimes denoted a spontaneous or nonce borrowing; see Sankoff, Poplack, & Vanniarajan, 1990). Depending on the sentence, the verb is inflected, as in the following two examples. There, the borrowing is adapted into French not only phonetically but also morphologically (note that there exists no French word “tagué”; it is, of course, the English guest word “tag”):

(4)	Oui, on a	tagged	ces cinq jupes.	CS: E /tʰægd/
(5)	Oui, on a	tagué	ces cinq jupes.	B: F /tage/
	“Yes, we have	tagged	these five skirts.”	

Since BIMOLA’s lexicon has its English verbs in the base form only, the model will not be able to recognize inflected forms such as “tagged”. In addition, the model does not currently account for the interesting but complicated process of morphologically adapting words from one language into another; so, neither will it recognize “tagué”. BIMOLA does, however, cope with all the guest words of the other examples: it can identify English words like “drop”, “tag”, and “peel”, be they used as code-switches (i.e. with guest language pronunciation, English /dɹɒp/, /tʰæg/, /pʰiɫ/) or be they employed as borrowings (i.e. with base language pronunciation, French /dʁɔp/, /tag/, /pil/). Vice versa, it can recognize French words code-switched (or borrowed) into English.

From its shared feature units, BIMOLA activates and inhibits phoneme units and word units in both its languages entirely independently yet in parallel. It does so just in the manner we have described in the general mechanisms (cf. phoneme activation and inhibition, word activation and inhibition) but separately within either of the two language networks: from the features via base language phonemes to the base language words, and from features via guest language phonemes to guest words. For BIMOLA, in contrast to FN5, we switched on the mechanism of top-down activation from words back to phonemes (McClelland & Elman, 1986; Elman & McClelland, 1988), with a small parameter value of 0.01; therefore, BIMOLA sends some activation from the base language words down to the base language phonemes as well as from the guest words down to the guest language phonemes. Lateral inhibition for guest language phonemes, as compared to lateral inhibition for base language phonemes, is attenuated by a parameter value of 4; that is, guest language phonemes compete among each other less severely than base language phonemes. All the guest language units (both phonemes and words) commence from lower resting levels than the base language units (phonemes and words), due to language mode and the guest language’s generally lesser language activation level (see above). Thus, guest words are usually activated less rapidly than base language words (and indeed, bilinguals take more time to access code-switches than base language words; Soares & Grosjean, 1984). Even if guest words begin to be active later, they do become active! With which speed a given guest word is activated in BIMOLA, and whether the word can finally be isolated or not, depends on a number of phenomena inherent to the guest word. As we will see later, in the evaluations, BIMOLA is able indeed to simulate several specific factors that were found to be involved in bilingual guest word recognition (Grosjean & Soares, 1986; Grosjean, 1988).

Box 11. To test guest words (i.e. code-switches or borrowings) in BIMOLA.

To test a code-switch in BIMOLA, just type the word into the "Word" field. Its transcription will appear and you are ready to run it in BIMOLA. In the case of cross-language homographs (e.g. the verb "gaze" exists both in English and French), you need to use the "Language" radio buttons.

To test a borrowing in BIMOLA (i.e. a guest word adapted phonologically to the base language), first do the same thing as for a code-switch and type the word into the "Word" field. As a result, BIMOLA will know that this word is the target. Remove the proposed transcription by clicking "Clear" next to the transcription. Now, depending on the direction of borrowing, open either the panel "E phon." (English) or "F phon." (French), and use the phoneme buttons that come into view to key in the borrowing's pronunciation as a series of phonemes. For example, to test the English word "toss" as a borrowing into French, type "toss" into the "Word" field, clear the English transcription (i.e. /tʰɒs−/), then open the "F phon." panel, and use the French phonemes to type a French pronunciation for "toss" (probably something like /tɔs−/). Always add the final silence input! You are now ready to run "toss" borrowed and adapted to French, in BIMOLA.

We should emphasize that there are not any inhibitory links whatsoever between BIMOLA's two languages, neither directly (e.g. from one set of words to the other set of words, if not from one set of phonemes to the other), nor indirectly via language nodes as used in the visual word recognition model BIA (Grainger & Dijkstra, 1992; van Heuven et al., 1998; Dijkstra & van Heuven, 1998, 2002). In BIMOLA, language activation is distributed over all the units (phonemes and words) of a language and hence a specific node dedicated to this purpose is not necessary. Since phonemes and words are organized in two language networks in BIMOLA, they also do not need a tag to which language they belong (another argument that was advanced for language nodes). Apart from the fact that there is no empirical evidence that language nodes exist and that we do not know how these nodes would be created when a new language is learned (Léwy & Grosjean, 2008), cross-language inhibitory links would actually be totally counterproductive for the spoken word recognition model BIMOLA. The recognition of guest words, which in BIMOLA always start from lower resting levels (especially low ones in the case of a monolingual mode), would be virtually impossible if base language word units would inhibit guest language word units (and vice versa). Borrowings would suffer even more (as compared to code-switches): they are presented to the model with the base language pronunciation, which naturally activates base language phonemes more than it activates guest language phonemes; with a cross-language inhibition (between words or, even worse, between phonemes), borrowings would have no chance at all to be recognized. The approach we advocate in BIMOLA, that is, the simultaneous activation of both languages (one language more than the other, but without them inhibiting each other), combined with a simultaneous inhibition within the two languages (in case of mismatching and competitive evidence in either language), seems to us to provide the best avenue for modeling spoken word recognition in the bilingual.

CHAPTER 6. EVALUATING FN5

Evaluation method and tools

In order to assess the validity of FN5, our model of monolingual spoken word recognition in French, we have followed a combination of two methodical approaches. On the one hand, we have determined the model's general recognition performance: How well is FN5 capable of recognizing an arbitrary word or sequence of words? In doing this type of evaluation, we have randomly selected a relatively large number of stimuli (single words or sequences of two words), have presented them to the model, one stimulus after another, and have so established the model's overall success rate in recognizing these stimuli. On the other hand, we have examined, in a number of evaluations of effects, FN5's ability to simulate specific psycholinguistic phenomena described in the literature: Does the model account, for example, for the well-known effect of word frequency? These are parametric studies with some independent and dependent variables (and possibly interactions), at a certain number of levels (usually two or three). We have selected the words or sequences of words to be tested at each level of a variable (and have, of course, strictly controlled for the other variables), have run these materials in the model, and have analyzed the results obtained in a classic manner (central tendency and dispersion, hypothesis test, analysis of variance, etc.).[32] As to the precise experimental effects that we have investigated, there are, as we will see, fundamental ones that are language independent (word frequency, length, and uniqueness point), and there are effects that are pertinent to spoken word recognition in French in particular (vowel duration, schwa deletion, and linking with and without liaison). We will report the FN5 evaluations in one section for isolated words and in another section for connected words, both times beginning with the general recognition study and then continuing on to the studies on specific effects (as shown in Table 6).

Table 6.
Outline of the FN5 evaluations

Evaluations on isolated words
General recognition
Effect 1: Word frequency
Effect 2: Word length
Effect 3: Word uniqueness point
Effect 4: Vowel duration
Evaluations on connected words
General recognition
Effect 5: Schwa deletion
Effect 6: Linking with and without liaison

Just before we get going, let us briefly introduce some helpful tools that have been used, and that will be referred to, in all the evaluations.

<u>*Lexical statistics database and search tool.*</u> The first of our evaluation tools, called “LexStats”, serves to find suitable test words for the parametric evaluations and to attend to control variables. It exists both in standard French and Swiss French versions, and is organized as a database in Excel format with 20,523 rows (one per word or pronunciation variant). We open this tool by simply clicking on the Excel file; then we apply one or several search filters so as to restrict the words to those corresponding to certain values or criteria of

[32] The statistical analyses were performed using the R language and environment (Ihaka & Gentleman, 1996; R Development Core Team, 2014).

interest (e.g. to all the words with a length of three phonemes and/or those that begin with the phonotactic pattern CV, etc.); in addition, we can sort the whole database by one or a combination of columns (e.g. first by word frequency and then by uniqueness point); we continue our search until locating the words we need; finally, we save our work and quit the tool.

LexStats contains basic properties (number of the word in the lexicon, number of the variant for a word, spelling, pronunciation, C/V pattern, length in terms of phonemes, and of syllables, gender, category, word frequency, and schwa preference where appropriate) and also includes attributes that we have calculated for each word, or pronunciation variant, in relation to FN5's whole lexicon: the uniqueness point (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1984), in number of phonemes as well as transformed to a percentage of word length[33]; both the size and the mean frequency of the neighborhood that can be formed by either substituting, adding, or deleting one phoneme at any position in the word (as defined by Luce et al., 1990); the number of homophones, etc. The algorithms required to compute all these values have been programmed by ourselves and are quite complicated because, for words with multiple variants, each variant got a separate row in the database and had to be compared with all variants of all the other words of FN5, excluding variants of the word itself. LexStats is available upon request, including user documentation, so as to encourage and facilitate further simulation studies with FN5.

<u>*Segmentation tool.*</u> This tool reveals the various ways in which a given sequence of phonemes can be divided up, in its entirety, into a sequence of words, disregarding syntax and semantics. To use this tool, we load a lexicon, enter a sequence of phonemes (by typing a word or sequence of words, or by choosing any sequence of phonemes freely), and then simply click on a button (labeled “Segm.”). A few examples of queries are presented in Table 7, along with the results for FN5's standard French lexicon. We observe that a sequence of phonemes sometimes correspond to one word (e.g. “lagune”), at other times can be segmented into two or more words (e.g. “la tour”), or may offer both possibilities (e.g. “labeur” and “la beurre”), despite the fact that these three examples have the same length and structure (/la/+CVC). Examining the fourth example (“la sauterie”), we see that certain sequences of phonemes offer a vast

[33] A value equal to 100% means that this point occurs on the word's last phoneme, and a value larger than 100% indicates that it is after word offset (the latter is the case, for example, for “jour” as there exist the words “journée”, “journal”, etc.).

Table 7.
Examples of using the segmentation tool

Query	*Results*
(1) Segment /lagyn/	lagune → 1 possibility
(2) Segment /latuʁ/	la + tour → 1 possibility
(3) Segment /labœʁ/	labeur la + beurre → 2 possibilities
(4) Segment /lasotʁi/	l' + as + au/haut/eau + tri l' + as + aut'/haute,haut~/hôte + ris/riz l' + assaut + tri la + sauterie la + c' + au/haut/eau + tri la + c' + aut'/haute,haut~/hôte + ris/riz la + sot/saut/sceau/seau + tri la + saute + ris/riz lasso + tri → 9 possibilities
(5) Segment /lasodʁi/	→ 0 possibility

number of alternatives of where to place word boundaries, even though the segmentation tool joins homophones (e.g. "au/haut/eau") into a single line of results. A primary cause for complexity are short words and contracted forms (such as "l' " and "c' "). The fifth example is obtained by changing one phoneme from the previous sequence of phonemes (/t/ becomes /d/); the new sequence of phonemes still is phonotactically legal, but it is no more a sequence of words in this lexicon.

Macro functionality. We have set up a number of macros, that is to say, small scripts containing sequences of commands (e.g. look up this word, input one or all of its phonemes, take a screenshot, write its isolation point to the log, and save the log to a file), which the model can automatically execute one after the other, and we have prepared lists of words, on which these macros operate. All the macros can be used time and again (also with new lists of words) and may be combined with other macros to form more complex macros, which set off a whole series of simulations in the model. We offer two interfaces for macros: a basic one, called “MacroShell”, which helps one to simply obtain the isolation point (IP) for each word or sequence of words in a user-editable list (cf. the bottom of Box 6, p. 68), and a richer interface, “MacroExpert”, allowing one to access and invoke the model’s full set of macros and lists, and to create new ones. There are “initialization macros”, serving to load a desired configuration (lexicon, parameter setting, etc.), at the beginning of a session; “run macros”, to launch a series of operations on a list of words (e.g. running in the model the recognition of single words or sequences of two words, running words either with or without schwa, etc.); “get macros”, used to obtain specific measures (such as the IP) of a word; and finally, “batch macros”, which let one carry out, without further intervention, one or several of the various evaluations (e.g. the evaluation on word frequency) in their entirety.

Without question, the macro functionality has been highly beneficial to us, since it has facilitated the whole testing process greatly, and has made it quick and reliable. It has let us redo evaluations after a change (while we were still developing the model), or has allowed us to easily run the same simulations under different conditions (e.g. standard vs. Swiss French lexicons, normal vs. fast speech rate, etc.). What is more, it enables any colleagues who wish to do so, to actually replicate some, or even all, of the evaluations that we will report, with our assurance that they will obtain the same results. For just this purpose, Appendix C1 provides a list of the main macros used, along with their role in the FN5 evaluations. As an aside, we point out that the automation approach by macros has, of course, not prevented us from also doing plenty of informal manual simulations (not reported here). Only by carefully examining how the word recognition process unfolds phoneme by phoneme, we were able to make sure that the model behaved as we intended, or find out (and then set right) a situation where it perhaps did not.

Isolated words

We will now describe the evaluations conducted on isolated (i.e. single) words. We begin by examining to what extent FN5 recognizes isolated words overall (general recognition), and we continue by studying how the model copes with four specific psycholinguistic effects (frequency, length, uniqueness point, and vowel duration).

<u>*General recognition.*</u> For this evaluation, a list of 1,000 words (nouns, determiners, and prenominal adjectives) was drawn at random from FN5's standard French lexicon ("FrenchPtitami.Lex"). Each test word appears only once. Using the segmentation tool described above, we found that 392 of the 1,000 words can be divided up phonemically into two or more words, in one way or another (cf. Table 7, Example 3, "labeur" → "la beurre"); the other 608 words cannot (cf. Table 7, Example 1, "lagune").

After loading the standard French lexicon and default set of parameters ("FN5.Diss.Set") into the model, the 1,000 words were presented to the model, one word at a time (using our macro functionality), to establish whether FN5 is able to recognize them. For each word, it counted as a success if the correct word was isolated[34] no later than the end of the silence segment that always followed the word (additional cycles were not allowed); it counted as a failure otherwise. In this way, the model's general recognition success rate (= number of successes ⁄ number of items × 100%) was computed. The entire test of the 1,000 words was carried out, separately, under the following two conditions:

A. The number of words being presented was unknown to the model.
B. The number of words being presented was known to be one, that is, an isolated word.

As we have described in the presentation of FN5's specific mechanisms, its PP (or position processor) can impose certain contextual constraints and can verify whether an activated word candidate fits those constraints. It was therefore just a question of whether or not we asked FN5 (i.e. the PP) to check the constraint of number of words. As we have explained, this is regulated by the parameter "Context: number of words", which was appropriately set to No for Condition A, and to Yes for Condition B.

We obtained the following results (also shown in tabular form in Table 8). In Condition A, the success rate was 91.3%, meaning that just 87 of the 1,000 words were not isolated by FN5. This is even better than we could expect since 392 words can be segmented into two or more words, as we mentioned above, and these words might have been activated by FN5 like that (i.e. as two or more words). We examined each of the 87 words that failed and we found that the

---

[34] We recall from p. 66f that a target word is said to be "isolated" at the moment in the simulation where it exceeds the response strength of all other words, homophones excepted, and from which on it retains this top rank until the end of the simulation.

main reason why FN5 made a mistake was indeed an unintended segmentation into several shorter words, typically words of a higher frequency (such as e.g. “remorqueur” → “remords + cœur”, “montant” → “mon + temps”, “inculture” → “un + culture”, “plâtrière” → “plat + tri + hier”, “salinité” → “sa + lit/lie + nid + thé/té”, etc.).[35] A secondary source of error was a dead end (or garden path, cf. Grosjean, 1980) into a beginning of a segmentation containing very frequent words (e.g. “un + part + donne” for “impardonnable”), out of which the model could not break out, either not at all, or not in time before the end of the silence. Under Condition B, the recognition success rate was 99.7%. Only 3 of the 1,000 words were not isolated, which shows that the PP mechanism responsible for constraining the number of words (and rejecting sequences like “remords + cœur” for “remorqueur”) was doing a good job. The three words that still were not isolated in Condition B are: “bienvenu” (due to an unlucky combination of schwa deletion, /bjɛ̃vny/, and mutual inhibition with homophonous “bienvenue”); “tain” (because “un” is highly activated and the homophones “teint/thym” are also present); as well as “plat” (here, at the end of silence, “la” is most activated but, with additional cycles, “plat” would win over it).

Evidently, FN5 is able to recognize isolated words very well. Once the number of words being presented (i.e. one word) is signaled to the model, as it is the case in Condition B, the general recognition success rate comes close to 100%. Condition B, the model’s default setting (adopted by the model unless the other setting is specified by the user), was maintained for the rest of the evaluations on isolated words, which deal with specific experimental effects.

Table 8.
FN5’s general recognition success rate for isolated words

Test condition		*Success rate*
No. of words was unknown	(A)	91.3%
No. of words was known	(B)	99.7%

35 Aside from these cases, and like TRACE and ARTWORD, FN5 shows a natural tendency to attach later arriving phonemes to the end of the current word and hence often prefers a long word (e.g. “caban” or “cormoran”) over several short words (“cas + banc”, “corps + mort + an”), even when the short words are more frequent.

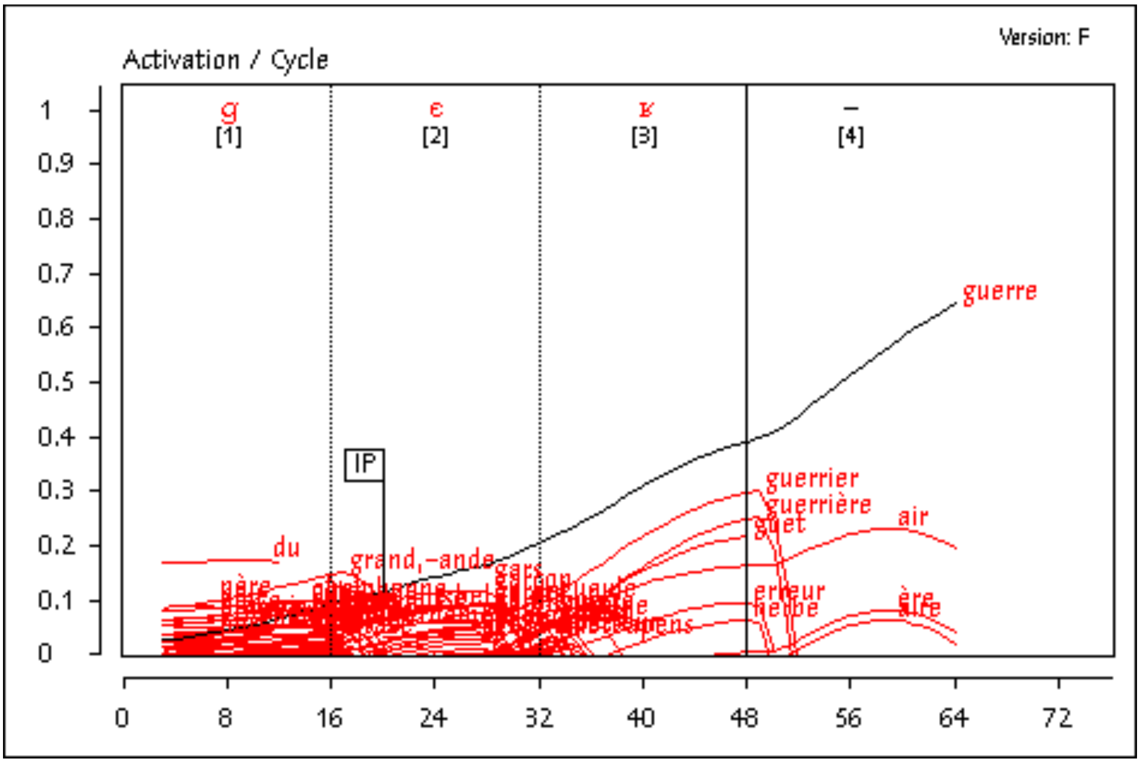
Frequency. The first effect we evaluated, the effect of word frequency, is a truly well-known and well-documented finding in over half a decade of literature on spoken word recognition: frequent words are recognized more quickly than infrequent words (Howes, 1957; Rosenzweig & Postman, 1957; Grosjean, 1980; Grosjean & Itzler, 1984; Connine et al., 1993; Walley et al., 1995; Metsala, 1997; Dahan et al., 2001; Cleland, Gaskell, Quinlan, & Tamminen, 2006; Dufour, Brunellière, & Frauenfelder, 2013).

With the help of our LexStats database, we chose thirty words of high frequency (value in the lexicon > 0.5, $M = 0.603$, $SD = 0.048$) and thirty words of low frequency (value in the lexicon < 0.2, $M = 0.087$, $SD = 0.046$), all of them nouns. The two groups of words were controlled for their length (in terms of number of phonemes and number of syllables), their uniqueness point, and also their neighborhood (both the size and the mean frequency). We ran these stimuli in the model and found that one low-frequency word ("daim") was not isolated by the model; it was therefore eliminated from the analysis. For all the other words, we determined the two IP measures described in the section on word isolation: IP@cycle (isolation point in absolute cycles) and %IP (isolation point in percentage of word length including silence). For both measures, there was an effect of frequency: high-frequency words are isolated by FN5 more quickly (i.e. need a smaller part of their unfolding input) than low-frequency words: the mean IP@cycle was 26.300 ($SD = 11.748$) for high-frequency words and 51.966 ($SD = 12.880$) for low-frequency words, $t(57) = -8.002$, $p < .001$; the %IP was on the average 40.017% ($SD = 18.788\%$) for high-frequency words and 75.356% ($SD = 9.669\%$) for low-frequency words, $t(57) = -9.036$, $p < .001$. Therefore, we conclude that the frequency effect is simulated by FN5.

Let us actually observe and demonstrate this effect in FN5 by showing two examples of word recognitions in the model, one above the other, and by comparing them graphically. Figure 12 visualizes the recognition of "guerre", a three-phoneme (one-syllable) word of high frequency (value in the lexicon = 0.553), and of "motte", a word with the same length but of low frequency (value in the lexicon = 0.120). In each diagram, activation levels of all the candidates proposed by the model are shown in function of simulation cycles; the target word ("guerre" and "motte", respectively), the IP of which is marked with a little flag, is drawn in black[36], the other candidates in red; phoneme and silence inputs (16 cycles each) are delineated vertically for easy reference. As one can see in the top diagram, the word "guerre" appears as a candidate already

---

[36] In white (on a gray background) in the simulation program.



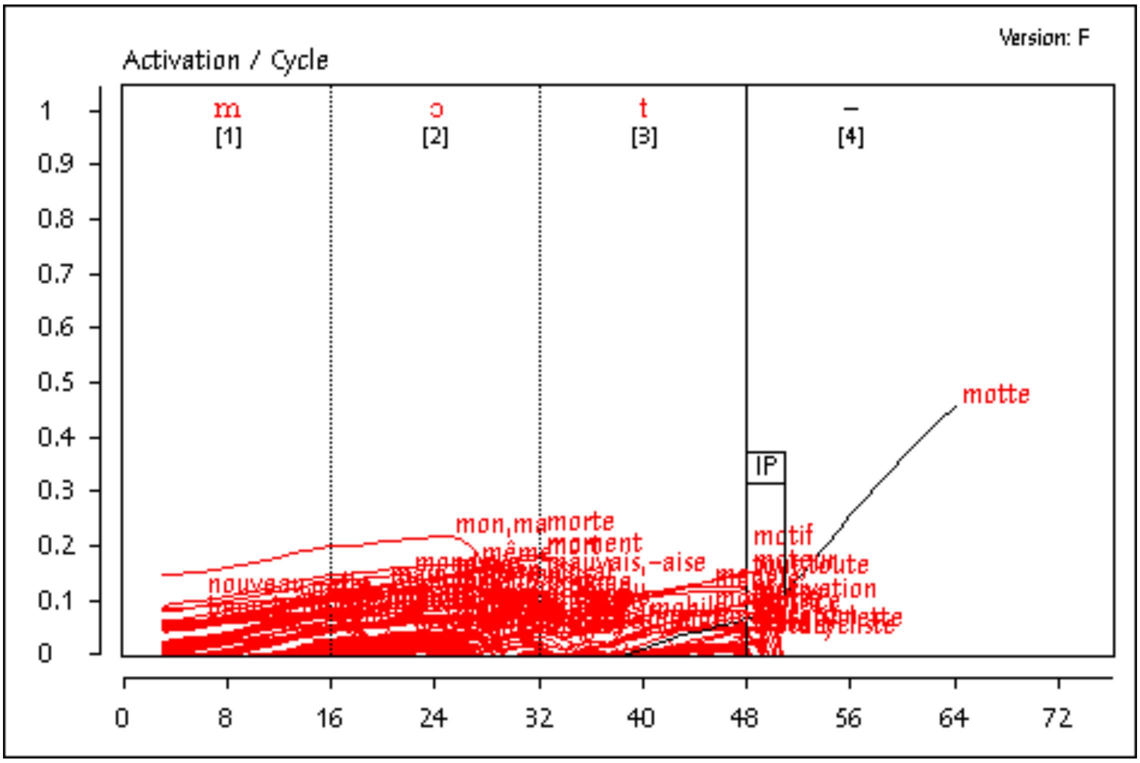


Figure 12. Simulating the recognition of “guerre” (/gɛʁ/), a high-frequency word, and of “motte” (/mɔt/), a low-frequency word.

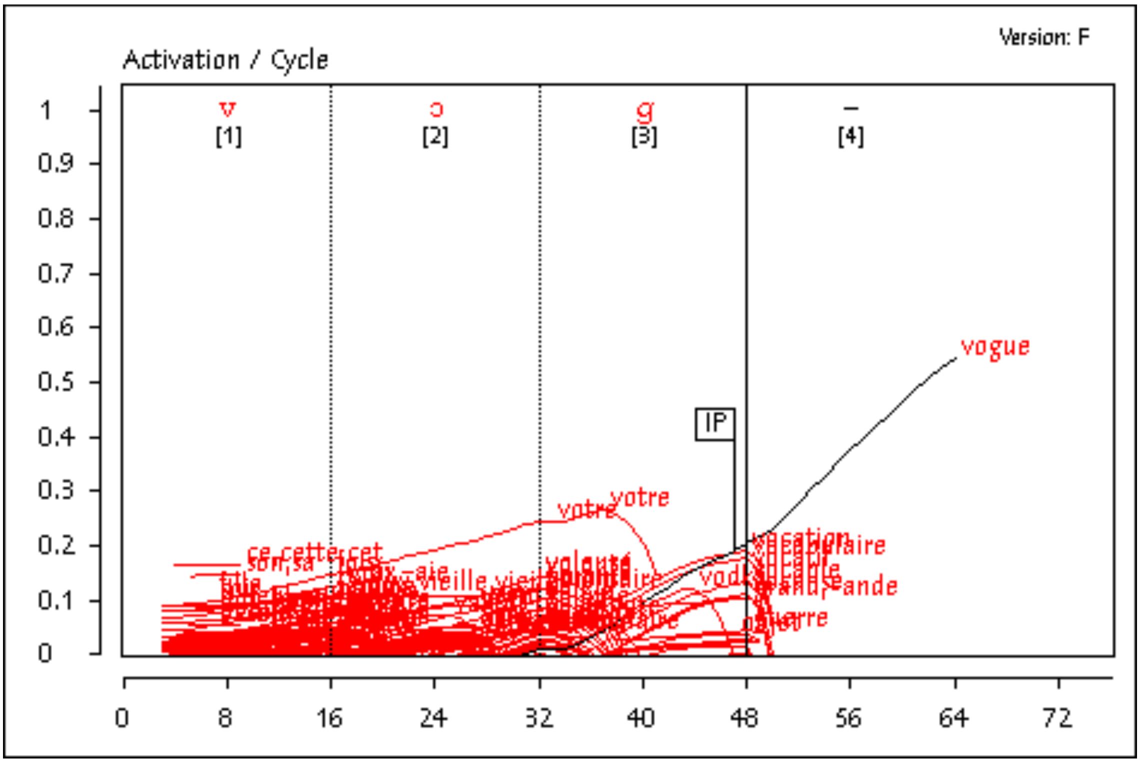
during the first phoneme and is isolated in the course of the second phoneme (IP@cycle = 20, %IP = 31.25%). By contrast, “motte”, in the bottom diagram, does not start to be activated before some point in the third phoneme and is isolated only during the silence after the word (IP@cycle = 51, %IP = 79.69%).

Length. Several researchers (e.g. Grosjean, 1980; Craig & Kim, 1990; Pitt & Samuel, 2006) have established that the processing of spoken words is also influenced by a factor of word length: short words (as expressed by the number of phonemes, or of syllables, they contain) are recognized at an earlier moment in time than long words.

To test the model for this effect, thirty short words and thirty long words (all nouns) were brought together. While the short words had three phonemes and were all monosyllabic, the long words contained six phonemes and were either disyllabic (23 words) or trisyllabic (7 words); between the two groups, the words were controlled for frequency, uniqueness point, and neighborhood. We tested these stimuli in the model: all of them were isolated by the model. Since the relative measure of %IP neutralizes, by definition, the length of a word (with the silence after the word included), we are here only interested in the absolute measure of IP@cycle. We obtained a mean IP@cycle of 42.600 (SD = 7.342) for short words and of 78.400 (SD = 18.997) for long words, $t(58) = -9.628$, $p < .001$. So, there was an effect of length: short words are isolated faster by this model than long words. We infer that the length effect is simulated successfully.

Figure 13 shows two examples of simulations, again side by side, one for a short word (“vogue”) and one for a long word (“coursier”). These two words have nearly the same frequency values (0.129 vs. 0.127) and both have a uniqueness point at their end. Nevertheless, the short word “vogue” (three phonemes, one syllable) is isolated much earlier, in term of cycles, than the long word “coursier” (six phonemes, two syllables): “vogue” is isolated at cycle 47 and “coursier” at cycle 83. The reason is that the longer word contains twice as many phonemes and hence obviously requires more processing cycles to be presented to the model, be it in full or in part, than the shorter word which has half as many phonemes. That the two words have a similar %IP (73.44% for “vogue” and 74.11% for “coursier”) does not surprise since the %IP measure reflects the isolation point normalized by word length. That is, both words are isolated at around 73% and 74% of their length plus the following silence.

Uniqueness point. The uniqueness point (UP) of a spoken word, as we may recall, refers to the earliest phoneme at which the word, when unfolded as



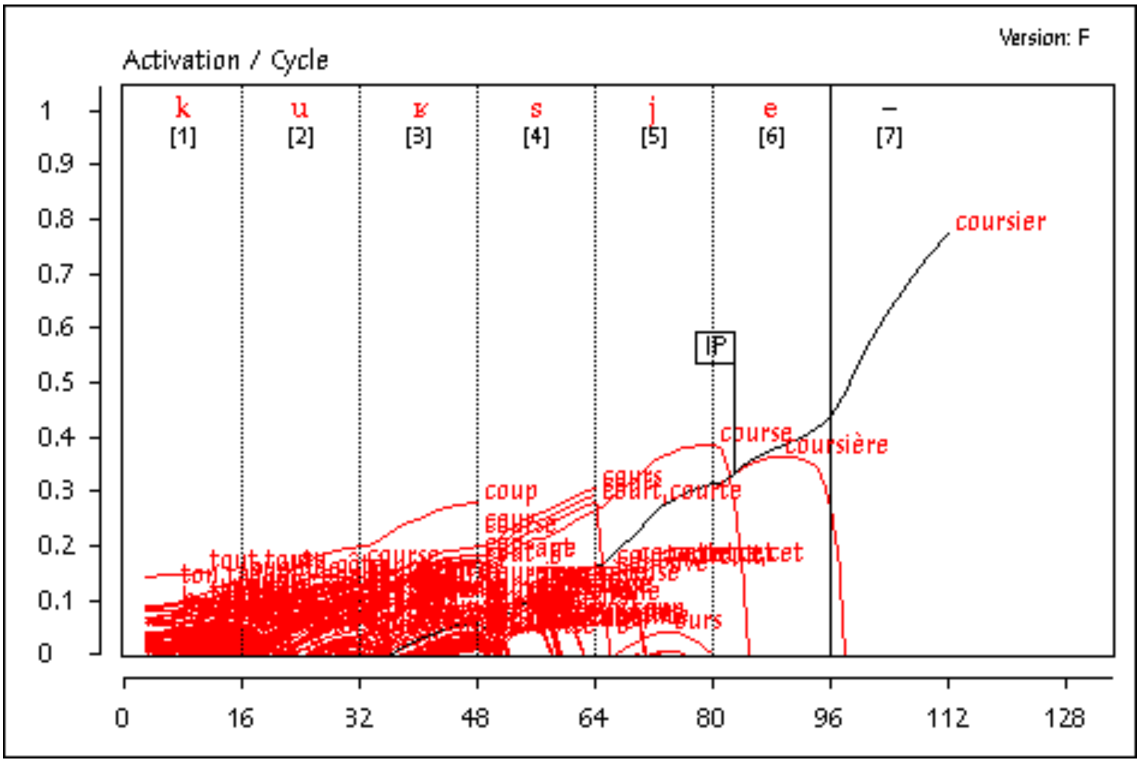


Figure 13. Simulating the recognition of “vogue” (/vɔg/), a short word, and of “coursier” (/kuʁsje/), a long word.

a sequence of phonemes from beginning to end (i.e. from its first phoneme to its last phoneme), begins to differ from all other words (homophones excepted) and becomes unique. Words that have a UP early in the word are recognized sooner than words that possess a UP late the word, as has been established experimentally time and again (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1984, 1987; Taft & Hambly, 1986; Radeau, Mousty, & Bertelson, 1989; Pitt & Samuel, 1995; O'Rourke & Holcomb, 2002; Henderson, Weighall, Brown, & Gaskell, 2013).

In contrast to the evaluations up to this point, we did not sort out our own materials for this evaluation but took the stimuli from the study by Radeau et al. (1989): 34 words that have an early UP and 34 words that have a late UP, once more all nouns. Whenever reusing someone else's materials for an evaluation with the model, one must verify first whether these materials keep their desired properties within the specific bounds of the model's lexicon (e.g. does a word that Radeau et al. used in their group of early UP also have an early UP in our lexicon?). We therefore made sure, basing ourselves on the values found in our LexStats database, that the two groups of stimuli indeed differ statistically as regards the tested variable, the UP (M = 58.429%, SD = 13.730%, for the early UP group; M = 90.465%, SD = 11.993%, for the late UP group, in percentage of word length), and that they do not differ statistically in respect of the control variables (frequency, length, and neighborhood); this was the case. We then presented the stimuli to the model, taking care to use pronunciation variants exactly as they are specified in Radeau et al.'s materials (e.g. "chemisier" with a schwa but "machinerie" without one). We obtained the following results: All the words were isolated; the IP@cycle had an average of 56.059 (SD = 8.198) for words with early UP and of 85.176 (SD = 21.071) for words with late UP, $t(66) = -7.510$, $p < .001$; the mean %IP was 44.082% (SD = 6.130%) for words with early UP and 65.718% (SD = 14.422%) for words with late UP, $t(66) = -8.050$, $p < .001$. Clearly, there was an effect of UP: words with an early UP are isolated by the model more rapidly than words with a late UP. That being so, the model succeeds in accounting for the UP effect.

Let us also here consider two examples of words, presented in Figure 14, one with an early UP and one with a late UP. The word "majuscule", which is pronounced /maʒyskyl/, has the UP on the /y/, the fourth of its eight phonemes (i.e. 50% into the word), since there exists no other word in our lexicon that begins with the sequence /maʒy/. As can be seen in the simulation at the top, this word is isolated by the model at the beginning of the /s/, the phoneme just following (IP@cycle = 65, %IP = 45.14%). The bottom example, "trésorier",

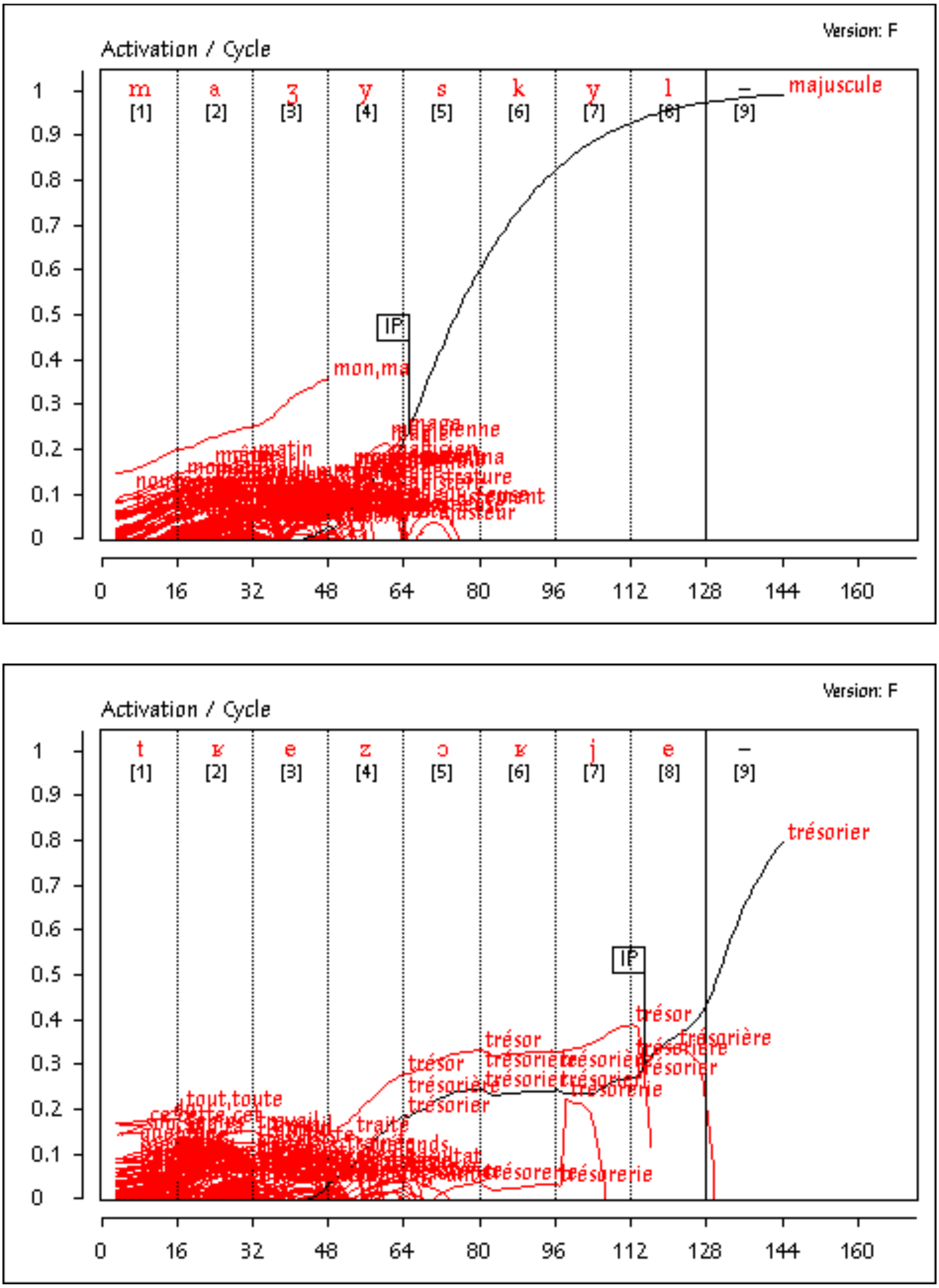


Figure 14. Simulating the recognition of “majuscule” (/maʒyskyl/), a word with an early UP, and of “trésorier” (/tʁezɔʁje/), a word with a late UP.

pronounced /tʁezɔʁje/, is of the same length of eight phonemes and of similar frequency as “majuscule” (0.084 vs. 0.077), but “trésorier” has its UP on the eighth phoneme, the final /e/; indeed, there are the words “trésor”, “trésorerie”, and “trésorière”, and thus one needs to know all the phonemes, up to the very last /e/ (not an /ɛ/ as in “trésorière”, /tʁezɔʁjɛʁ/), before one can tell whether “trésorier” is meant. We observe it in the simulation: “trésorier” is isolated during cycle 115 (%IP = 79.86%), far later than “majuscule”, despite the fact that the two stimulus words become active at roughly the same time (both black curves begin at cycle 43).

The role and importance of the UP in spoken word recognition has been critically challenged by Radeau, Morais, Mousty, and Bertelson (2000), who found, in a study with French-speaking participants and using just the material of Radeau et al. (1989), an effect of UP at two slower speech rates but none at a faster speech rate. They concluded that the UP effect is nothing but a strategy and that the finding of a UP effect should be regarded as a theoretical mistake. There is an alternative interpretation, which we prefer: The UP effect is an effect that materializes when a spoken word is being presented strictly sequentially from its beginning to its end (as it is the case at slower speech rates) but that disappears when parts of the word are being presented rather simultaneously (as it happens, at least to some degree, at faster speech rates). Or, as Mattys (1997) has put it, with slower speech, the recognition system has the necessary time to uncover the sequential build-up of a word (there results a UP effect); with faster speech, the sequentiality in the word is still there but the recognition system has no time to reflect it in its activation state (therefore the UP effect is masked). Interestingly, whereas Radeau et al. (2000) did not offer any model, we can with both our models, be it with FN5 or be it with BIMOLA, reduce the importance of the UP effect by simply speeding up the input to the models. As described earlier, the speech rate and hence the presentation mode (sequential vs. simultaneous) are accounted for in our models by varying one parameter, the phoneme input delay (cf. Figure 10, p. 58). When it is set to 16 cycles (i.e. to the default setting, which represents the normal speech rate and sequential presentation mode), phoneme inputs arrive in the model successively, one after another, and each can be processed by the model individually, and without haste (for 16 cycles), before the next phoneme input turns up; this causes the UP effect. When however the parameter is set to 3 cycles (which corresponds to a fast speech rate and partially simultaneous presentation mode), phoneme inputs enter the model much more quickly (every 3 cycles) and with some considerable overlap (because they still extend for 16 cycles); the model can devote only little time to each single phoneme input and needs instead to process several of them together at the same time; as a result, the UP effect cannot manifest itself enough.

All evaluations so far, including the one on the UP, used the phoneme input delay of 16 cycles. We now applied the phoneme input delay of 3 cycles and reran our UP evaluation (still on the stimuli of Radeau et al., 1989, 2000), predicting, with this fast speech rate, an absence of a UP effect. We had the following results: The words were still all isolated; now the IP@cycle had an average of 21.853 (SD = 1.617) for words with early UP and of 22.206 (SD = 2.783) for words with late UP, $t(66) = -0.639$, NS; and the mean %IP was now 59.371% (SD = 5.709%) for words with early UP and 59.604% (SD = 6.786%) for words with late UP, $t(66) = -0.153$, NS. For both measures, evidently, the difference of the means, which was quite large with the normal speech rate (in the preceding evaluation), became very small and statistically insignificant in the case of the fast speech rate. So the UP effect, which was clearly present for the normal speech rate, practically disappeared for the fast speech rate, just as we expected. In Figure 15, “majuscule” and “trésorier”, the same two examples of words as before, are shown again, but now when run at the fast speech rate, that is, phoneme inputs arrive every third cycle (e.g. the /m/ of “majuscule” at cycle 0, the /a/ at cycle 3, the /ʒ/ at cycle 6, etc.) and stretch beyond other phoneme inputs (e.g. /m/ extends from 0 to 16 and overlaps, in part, with /a/, /ʒ/, /y/, /s/, and /k/). Even though “majuscule” is an early-UP word and “trésorier” is a late-UP word, we observe that the two simulation graphs hardly differ in their general shape (cf. Figure 14 where the two graphs contrasted strikingly). In particular, we find that the advantage of “majuscule” compared to “trésorier” (which was, for the normal speech rate, of 115−65 = 50 cycles) has decreased, for the fast speech rate, to a single cycle: “majuscule”: IP@cycle = 23, %IP = 23 ⁄ [3 × 8 + 16] = 57.50%; “trésorier”: IP@cycle = 24, %IP = 24 ⁄ [3 × 8 + 16] = 60.0%. All this confirms our alternative interpretation and explanation, that is, that the UP effect is indeed an effect (and not an artifact, as argued by Radeau et al., 2000) but one that is sensitive to speech rate and presentation mode. We used the normal speech rate (i.e. sequential presentation mode) in the rest of the FN5 evaluations, and likewise in all of the BIMOLA evaluations.

Vowel duration. Since FN5 deals not only with standard French but also with Swiss French (the variety of French spoken in Switzerland), the evaluations of course also included a section on examining whether FN5 successfully differentiates between these two versions of French. One major

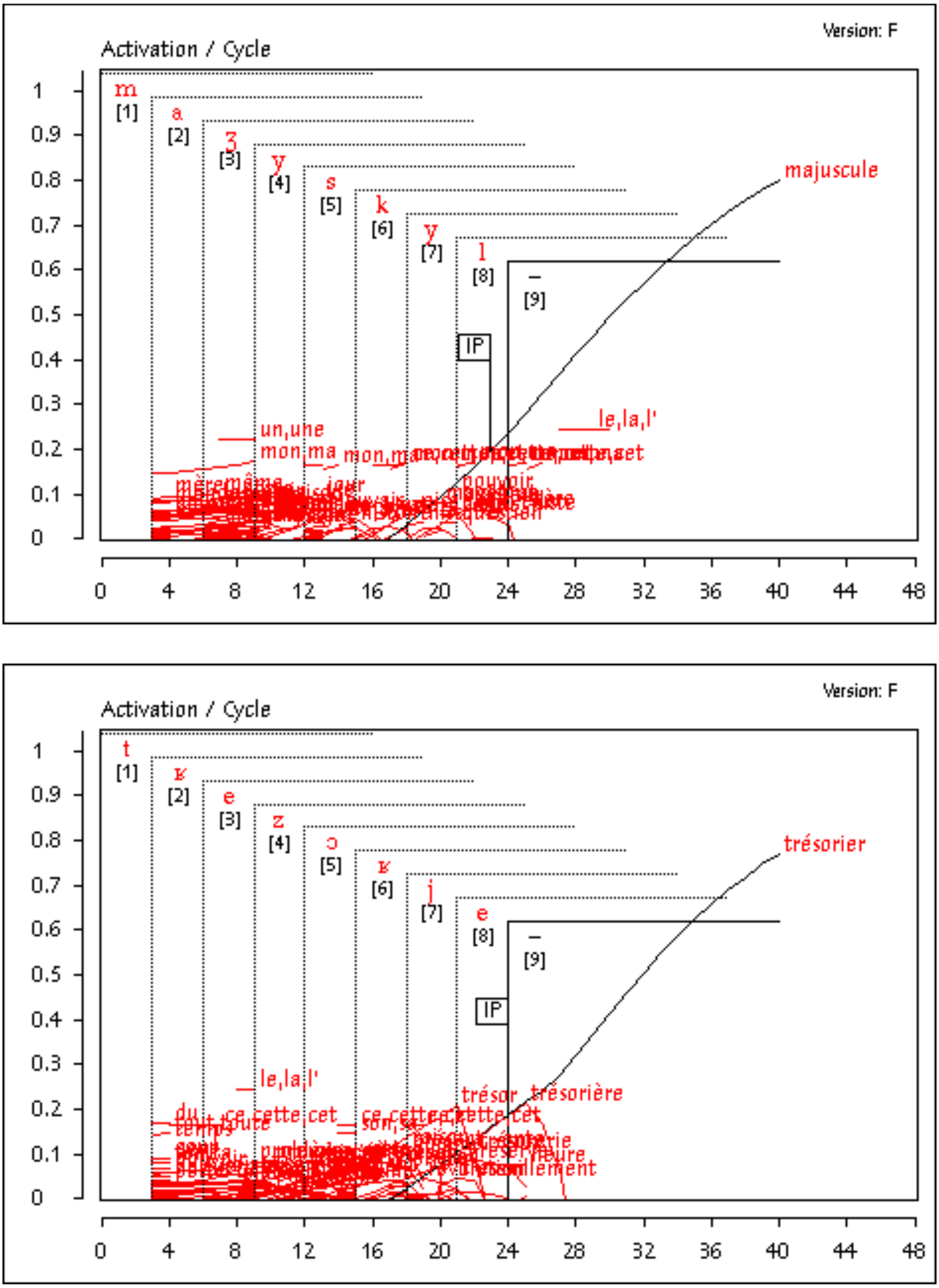


Figure 15. Simulating the recognition of the same two words as in the previous figure, but this time at a fast speech rate.

phonological difference is, as we pointed out, that long final vowels are absent in standard French but present in Swiss French (Métral, 1977; Knecht, 1985; Grosjean et al., 2007). For example, the words “mordu” and “mordue” are, in standard French, both pronounced /mɔʁdy/ and hence are homophones. By contrast, in Swiss French, “mordu” ends in a short vowel (/y/) but “mordue” in a long vowel (/yː/), that is, the vowel duration is distinctive.

To examine the model on the factor of vowel duration in Swiss French, we used thirty pairs of words of the type “mordu/mordue” (always one word of a pair with a short final vowel and the other with a long final vowel), which we took from Grosjean et al. (2007); there are nouns, adjectives, and verbal participles. A special lexicon file, containing the 60 words in question, along with subjective frequency values (drawn from Grosjean et al. and converted to the model’s own frequency scale), was loaded just for this simulation in addition to the normal lexicon; the words already present were replaced. We analyzed, as always, the control variables and found that the words ending in a short vowel have higher frequency values (M = 0.295, SD = 0.088) than the words terminating in a long vowel (M = 0.257, SD = 0.081, paired $t(29) = 4.139$, $p < .001$). (This also holds true for neighborhood size, but at the $p < .01$ level.) We anticipated therefore that the model would isolate short vowel words slightly faster than long vowel words, just as Grosjean et al. had found in their experiment with Swiss French participants.

In Figure 16, we first show the case of “mordu/mordue” for the standard French version of the model (i.e. the pronunciations in the lexicon, the set of units on the phoneme level, and the input to the model, /mɔʁdy/, were all in standard French). Since “mordu” and “mordue” have the same pronunciation, both words of this pair are activated in the simulation and, as their frequency values are very close (“mordu”: 0.259, “mordue”: 0.255), they almost share one activation curve (IP@cycle = 67, %IP = 69.79%). This contrasts sharply with what happens in the Swiss French version of the model (with the phoneme repertoire and word pronunciations now in Swiss French), which is shown in Figure 17 and drawn in green. In both simulations (“mordu” /mɔʁdy/ at the top, “mordue” /mɔʁdyː/ at the bottom), we see that the two words of the pair split up during the final vowel; each time, only the correct member of the pair continues to rise while the other one falls rather quickly. Additionally, we observe that “mordu” is isolated faster (IP@cycle = 67, %IP = 69.79%) than “mordue” (IP@cycle = 75, %IP = 78.13%). This comes to pass as “mordu” has a slightly higher frequency value than “mordue” (0.239 vs. 0.217). When we ran the remaining stimuli in the model, we found that the model correctly identified them all, both the words with a final short vowel and the words with a final long vowel. There was an effect of vowel duration: short vowel words were isolated more rapidly than long vowel words: the average IP@cycle was 52.067 (SD = 14.007) vs. 55.967 (SD = 13.795), paired $t(29) = -2.507$, $p < .01$; and the mean %IP was 71.156% (SD = 10.553%) vs. 76.233% (SD = 6.970%), paired $t(29) =$ -2.665, $p < .01$. It means that the model simulates vowel duration as it should. The effect we found is relatively small (3.9 cycles) and is most likely caused by the frequency value being higher for the words with a short vowel. Another factor at play could be the neighborhood size (larger for words with short vowel); it is a sign of the frequency of the vowels themselves (short vowels are generally more frequent).

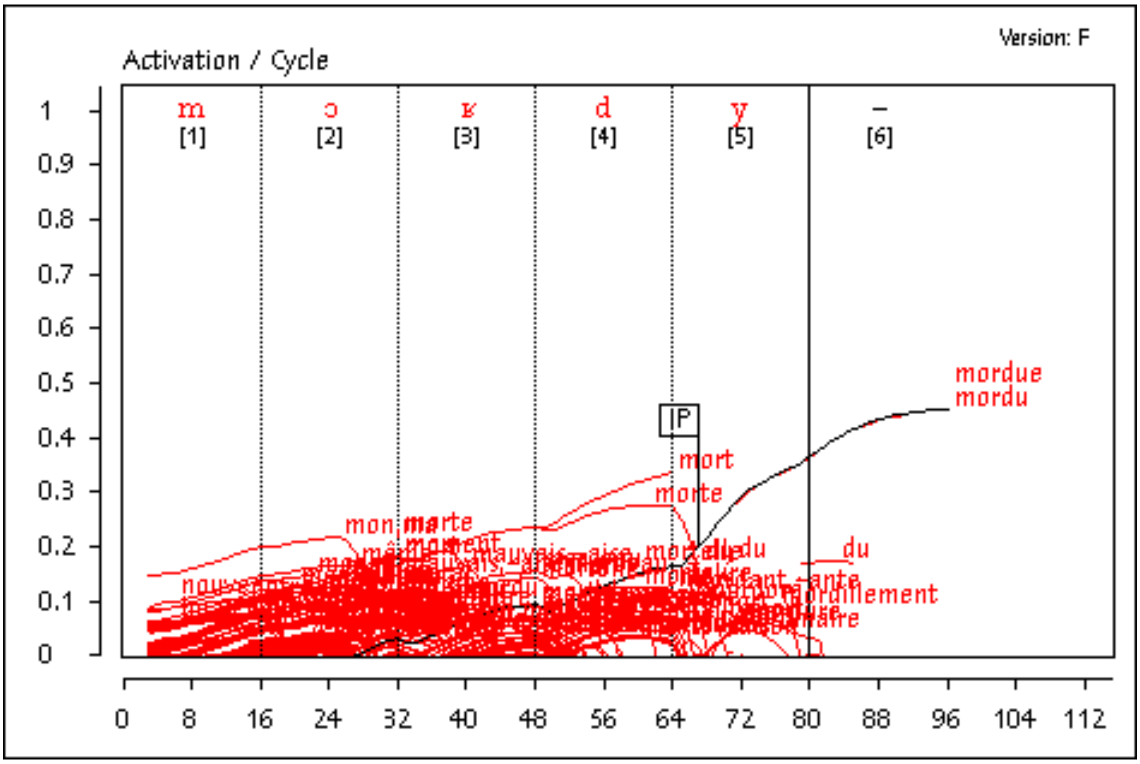


Figure 16. Simulating the recognition of “mordu” (/mɔʁdy/), homophonous with “mordue” in standard French.

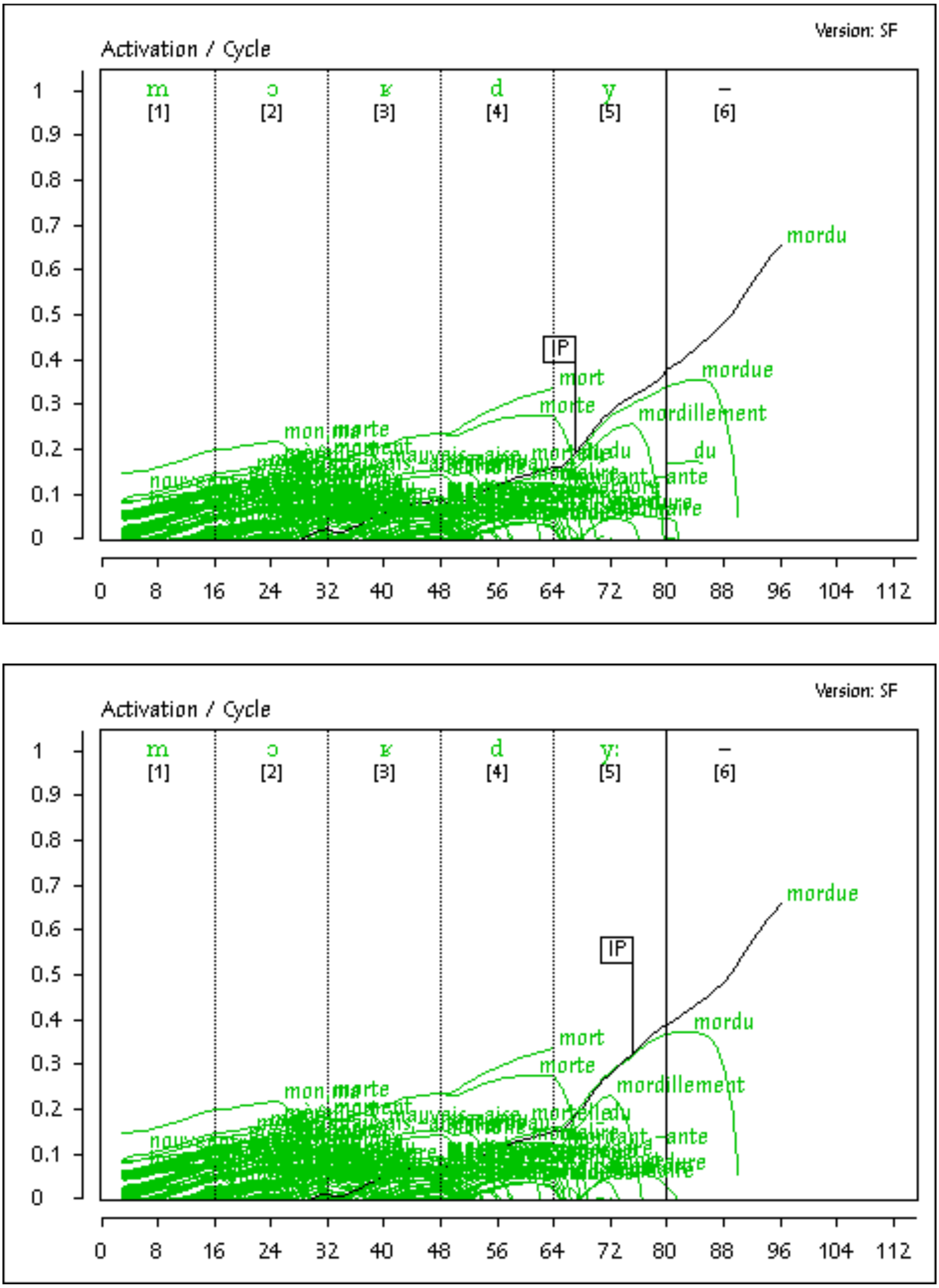


Figure 17. Simulating the recognition of “mordu” (/mɔʁdy/) and “mordue” (/mɔʁdyː/), which are non-homophonous in Swiss French.

Connected words

We now turn to the evaluations carried out on connected words. Just as we have proceeded for the isolated words, we start by assessing FN5's general recognition capacity, this time on connected words (i.e. sequences of words). We subsequently investigate two sets of specific effects related to connected words (deletion of schwa and linking with and without liaison).

General recognition. A list of 1,000 sequences composed of two words (either determiner + noun, or prenominal adjective + noun) was generated by random selection from FN5's standard French lexicon. Each test sequence is unique, though some words turn up in several sequences of two words, as chance would have it. All the sequences of words are syntactically well-formed, that is, gender agreement was observed and liaison was used where it applies; no attention was paid, of course, to semantics. Using the segmentation tool, we found out that 431 of the 1,000 sequences of two words have an unambiguous segmentation status: their phonemic sequence can be divided in one way only (cf. Table 7, a few pages back, Example 2, "la + tour"). By contrast, the other 569 sequences of two words have an ambiguous segmentation status: their phonemic sequence can be cut up in more than one way (cf. Table 7, Examples 3 and 4), sometimes even into two words in several ways (e.g. /lasotʁi/ → "la + sauterie" or "lasso + tri").

The 1,000 two-word sequences were tested one after the other. For each sequence, it counted as a success if both words of the sequence were isolated by FN5 at the latest by the end of the silence segment that always followed the sequence (additional cycles were not permitted). Thus, the general recognition success rate (= number of successes ⁄ number of sequences × 100%) was calculated. Sequences were presented to the model either as two connected words, that is, as a single sequence of phonemes (e.g. /gʁɑ̃taʒɔ̃/ for "grand ajonc"), or separated by a pause, that is, with a silence segment between the two words (/gʁɑ̃t−aʒɔ̃/). There is no boundary between words in the former case, while in the latter case, the boundary is explicitly marked. For either type of sequences, we ran the complete evaluation (of 1,000 sequences) separately under the following three test conditions:

A. The number of words being presented was unknown to the model.

B. The number of words being presented was known to be two, that is, a sequence of two words; the initial word of the sequence could come from any lexical category.

C. The number of words being presented was known to be two, and the initial word of the sequence had to be a determiner or prenominal adjective.

These test conditions are realized by appropriately setting the two parameters that are concerned with contextual constraints in the position processor (PP), “Context: number of words” and “Context: prenominal word at beginning”, that is, No/No for A, Yes/No for B, and Yes/Yes for C.

Table 9 summarizes the results of the six runs all in all (3 test conditions × 2 types of sequences). The row for Condition A conveys that 164 sequences of two connected words (and 137 sequences of two words separated by a pause) were not isolated. This outcome beats our expectations since 569 sequences have an ambiguous segmentation status and could have caused considerable difficulties (the model did not know that two words were being presented and was allowed to activate a single word or more than the two).[37]

Table 9.
FN5’s general recognition success rate for two-word sequences

		Type of sequence	
Test condition		*2 connected words*	*2 words separated by a pause*
No. of words was unknown	(A)	83.6%	86.3%
No. of words was known			
Any word at beginning	(B)	99.3%	99.1%
Prenominal word at beginning	(C)	99.9%	99.9%

[37] Most of the ambiguous two-word sequences correspond to sequences of three or more very short words (e.g. “simple allumeur” → “saint/sein + plat + l’ + humeur”, “certaine rumeur” → “certaine + rhume + heure”). As we already discussed, FN5 has an inclination to activate the longer words (“simple” and “rumeur”) more than the shorter words.

For this condition alone, the presence of a pause separating the two words, and thus clearly marking the boundary between the words, was of help to the model, but not as much as one might perhaps assume. Under Condition B, only 7 sequences of two connected words (and 9 sequences of two words separated by a pause) were not isolated, which is an almost perfect result, already. The model made a slip, for example, in coming up with “ville + biquette” (two nouns) instead of “vile + biquette” (adjective + noun). Finally, in Condition C, just one sequence of two words (be it with pause or without) was not isolated: “haute coagulation”. This happened for the sole reason that the model proposed the contracted form “aut’ ” (/ot/) of “autre”, in place of “haute”, which is pronounced the same way. All the other 999 sequences of two words were isolated correctly.

Overall, the model has a remarkable ability, so it appears, to recognize sequences of two words. When the number of words being presented (i.e. two words) and the lexical category of the initial word (determiner or adjective) are indicated to the model, as it is the case in Condition C, the general recognition success rate verges upon 100%. Moreover, the presence or absence of a pause (i.e. an explicit indicator of a boundary) between words is not crucial for the model. FN5 finds the word boundaries by itself, and it does so simply in the course of recognizing the words (as a by-product, so to speak). In consequence of these results, from here on, we will maintain Condition C (which corresponds to FN5’s default setting) and will always use sequences of connected words.

Schwa deletion. Words that contain a schwa are represented in FN5, as we remind, as single entities with multiple pronunciation variants (i.e. one variant with and one variant without schwa, or more variants in the case of a word with several schwas). Preference values, one per pronunciation variant (from Racine, 2008), are stored within these words, and are used to implement a variant frequency (cf. Connine & Pinnow, 2006; Bürki & Frauenfelder, 2012). All the words containing a schwa that happened to be a part of our general recognition studies were tested on their preferred pronunciation (the one with the highest preference value), whatever that pronunciation might be. But now, in this evaluation, we tested FN5 on the non-preferred pronunciation as well. We first assessed whether FN5 can show the (basic) effect of schwa deletion, namely that words are recognized less quickly with schwa deletion than without schwa deletion (Racine & Grosjean, 2000). We then examined if FN5 is able to differentiate that effect according to types of schwa words, that is, depending on whether the schwa deletion for a word is mandatory, optional, or prohibited (Racine & Grosjean, 2005).

The two-word sequences of Racine and Grosjean (2000) were made up of a determiner (“mon, ma”, “ton, ta”, “son”, or “la”) followed by a noun with an optional schwa, such as “son genou”: /sɔ̃ʒ(ə)nu/. Of their 16 sequences, we had to omit “ma jeunesse” because “jeunesse” is stored in FN5’s lexicon with an /œ/, not with a schwa; thus we used 15 sequences. We tested each sequence twice in FN5, a first time without schwa deletion (i.e. the /ə/ was present, such as in /sɔ̃ʒənu/) and a second time with schwa deletion (i.e. the /ə/ was absent, as in /sɔ̃ʒnu/). We had the following results. We found that the sequence “la petite”, in the condition of with schwa deletion (/laptit/), was not isolated in time; as we did paired comparisons between the two conditions, we removed this sequence from the analysis entirely.[38] All the other 14 sequences were isolated by FN5 in both conditions. The nouns’ mean IP@cycle (calculated from the beginning of each noun) was 43.571 (SD = 13.501) in the condition of without schwa deletion and 59.500 (SD = 7.198) in the condition of with schwa deletion, paired $t(13) = -5.055$, $p < .001$; their %IP was on the average 46.412% (SD = 15.898%) for without schwa deletion and 76.525% (SD = 12.760%) for with schwa deletion, paired $t(13) = -8.247$, $p < .001$. Clearly, there was an effect of schwa deletion: the words are isolated faster without schwa deletion than with schwa deletion. We should mention that when we verified the various control variables (what we did, as always), we found that the nouns’ versions without schwa deletion have a smaller neighborhood than the nouns’ versions with schwa deletion ($p < .001$ for neighborhood size and $p < .01$ for neighborhood frequency). This is in the nature of the phenomenon of schwa deletion and cannot be avoided.

Figure 18 visualizes the recognition of the sequence “ta semaine”, once without schwa deletion (shown at the top)[39] and once with schwa deletion (at the bottom). During the first three phonemes (i.e. /tas/), the same candidates are activated, and the determiner “ton, ta” is isolated (IP@cycle = 26, %IP = 54.17%), in either of the two versions. The divergence begins, as one would

---

[38] For “la petite” with schwa deletion (/laptit/), the sequence “l’aptitude” (/laptityd/) stays active for a long while and therefore strongly inhibits “la petite”. Even though “aptitude” is deactivated during the final silence that always follows our stimuli, “petite” does not get enough time then to change its activation level from negative to positive before the end of the simulation; it would need some additional cycles after the silence to do so and to be isolated by FN5.

[39] To increase legibility of the top diagram, we stopped that simulation just five cycles before its normal end. At the very end of the simulation, the word labels of “ta” and “semaine” are superimposed and hard to read.

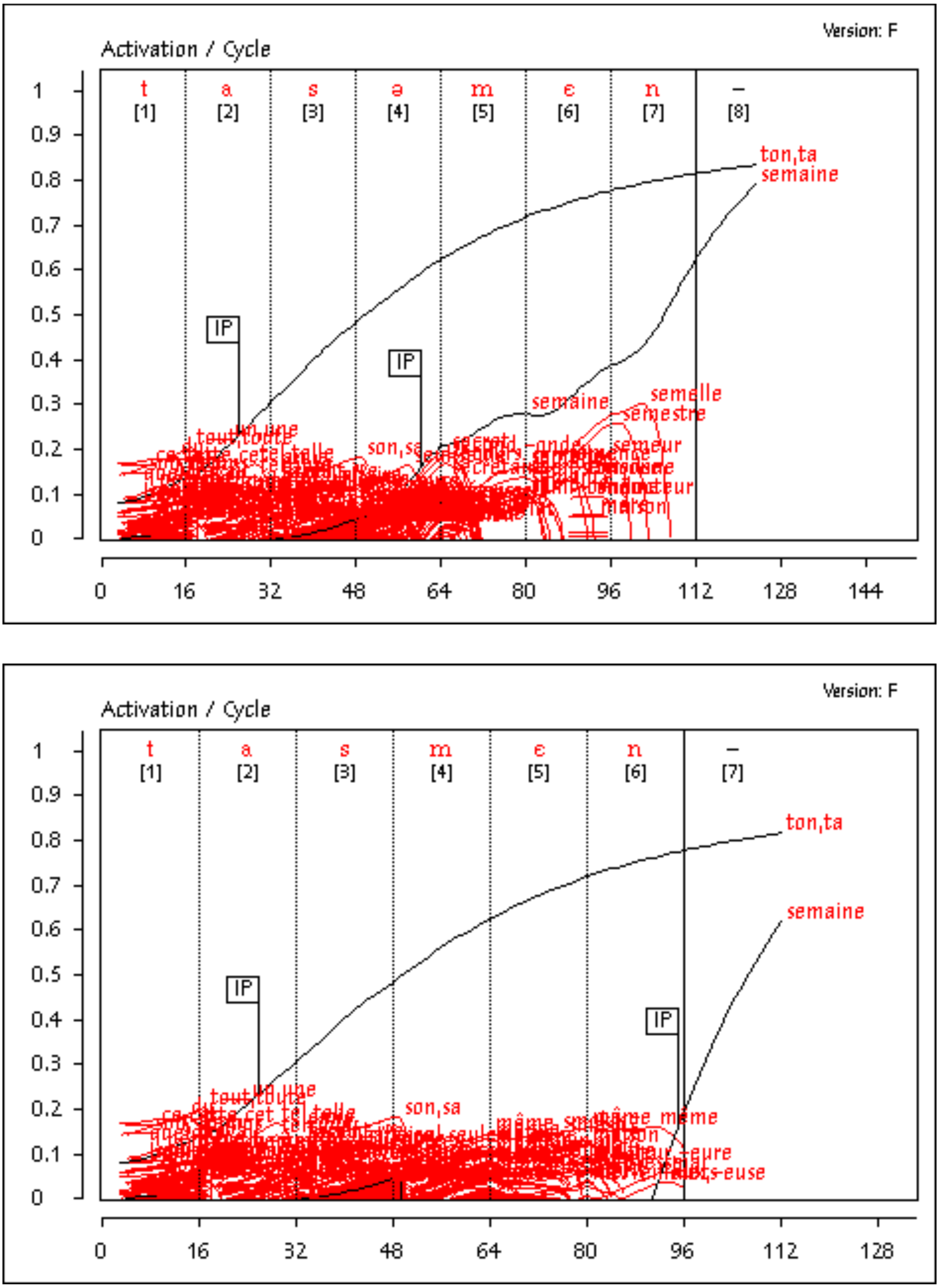


Figure 18. Simulating the recognition of “ta semaine”, without schwa deletion (/tasəmɛn/, top diagram) and with schwa deletion (/tasmɛn/, bottom diagram).

expect, at the fourth phoneme (/ə/ for the version without schwa deletion and /m/ for the version with schwa deletion). For the version without deletion, the word “semaine” is isolated during the schwa (IP@cycle = 60, IP@cycle Word = 28, %IP = 29.17%). By contrast, for the version with deletion, the /m/ fosters candidates starting with /m/ (e.g. “même”, “mère”, “maison”, “main”, etc.); the word “semaine” does not become active before its last phoneme (the /n/) and is isolated near that phoneme’s end (IP@cycle = 95, IP@cycle Word = 63, %IP = 78.75%), that is, a whole 35 cycles later than in the case without deletion. The two pronunciation variants of “semaine”, /səmɛn/ and /smɛn/, have, by the way, the same uniqueness point (on the /n/); that they differ in their neighborhood size (1 for /səmɛn/ vs. 4 for /smɛn/) is irrelevant since the additional neighbors of /smɛn/ (“scène”, “cène”, “hymen”) do not have any impact, and neither does the one neighbor (“semelle”) that is common to the two variants.

Let us turn to the study of Racine and Grosjean (2005), who put together 48 two-word sequences, comprising a definite article (“le, la”) and a noun with a schwa; for 16 of the nouns, the schwa deletion is mandatory (e.g. “la batterie”), for 16 nouns, the schwa deletion is optional (e.g. “le velours”), and for 16 nouns, the schwa deletion is prohibited (e.g. “le parmesan”). As Racine and Grosjean did their experiments in Neuchâtel, we ran FN5 for this evaluation in the Swiss French version: its schwa preference values differ from the standard French version, for some words just a little, for other words greatly. Again, we tested all the sequences in FN5 two times, once pronounced without schwa deletion and once pronounced with schwa deletion. In either condition, all 48 sequences were isolated by FN5. The nouns’ %IP was examined in an analysis of variance, with one within factor (with or without schwa deletion) and one between factor (word type: mandatory, optional, or prohibited schwa deletion). A main effect was found for the factor of schwa deletion, $F(1, 45) = 36.030$, $MSE = 98.378$, $p < .001$; the %IP was on the average 56.225% (SD = 12.498%) for without schwa deletion and 68.378% (SD = 13.987%) for with schwa deletion. (This repeats the effect we showed above for Racine and Grosjean’s (2000) stimuli.) No main effect was found for the factor of word type, $F(2, 45) = 1.264$, MSE = 239.486, *NS*; the mean %IP was 58.913% (SD = 11.510%) for the nouns with mandatory deletion, 63.070% (SD = 17.454%) for those with optional deletion, and 64.920% (SD = 13.823%) for those with prohibited deletion. There was an interaction between the two factors, $F(2, 45) = 3.692$, $MSE = 98.378$, $p < .05$; mandatory/without: 56.508% (SD = 10.847%), mandatory/with: 61.319% (SD = 11.992%), optional/without: 56.273% (SD = 16.401%), optional/with: 69.867% (SD = 16.189%), prohibited/without: 55.893% (SD = 10.127%), prohibited/with:

73.947% (*SD* = 10.886%). Post-hoc analyses revealed a significant difference between with and without schwa deletion for the nouns with optional deletion ($p < .05$) and for the nouns with prohibited deletion ($p < .001$) but not for those with mandatory deletion ($p = .056$). In the condition of with schwa deletion, the difference is significant between nouns with mandatory vs. prohibited deletion ($p < .01$) but not between the other word types (mandatory vs. optional deletion and optional vs. prohibited deletion). In the condition of without schwa deletion, there are no differences among word types; that is, when the schwa is present, words are isolated with the same speed whether the schwa deletion would be mandatory, optional, or prohibited. All of this matches the analytical results of Racine and Grosjean (2005): they too had a main effect for schwa deletion, no main effect for word type, and an interaction between the factors, and they too found, in their post-hoc analyses by word type, that the difference between with and without schwa deletion was largest for the words with prohibited deletion, somewhat smaller (but still significant) for the words with optional deletion, and smallest (and therefore non-significant) for the words with mandatory deletion.

To allow a visual comparison, Figure 19 shows our results in the upper panel and those of Racine and Grosjean's (2005) two perception experiments (one used the word repetition paradigm, the other used the lexical decision paradigm) in the two lower panels. Means, by word type (mandatory, optional, or prohibited schwa deletion) and by condition (with or without schwa deletion), are depicted in an interaction plot (see Keppel, 1991). The experiments were reported with a measurement of ratio of the participants' reaction time (in ms), divided by sequence length (in ms), which is comparable with our measurement of %IP, that is, isolation point (in cycles), divided by full word length (in cycles), even though we multiply by 100%. As one can see in the figure, the model's and the experiments' results correspond extremely well. There is one visual difference: the circle of mandatory/without is a bit too low for FN5 and should be a little higher up and closer to the triangle of mandatory/with. We could adjust this with the help of a small inhibitory effect (i.e. a bias < 0.5, see p. 65) for the former pronunciation variant. But, as we mentioned, even now, the difference between mandatory/without and mandatory/with is non-significant, statistically. So, there is no pressing need for a change.

Linking with and without liaison. In French, a spoken word ending in a consonant is normally linked to the beginning of the next word; the linking can be with or without liaison (*l'enchaînement avec ou sans liaison*; Encrevé, 1988; Gaskell et al., 2002; Spinelli et al., 2003; Yersin-Besson & Grosjean, 1996). In

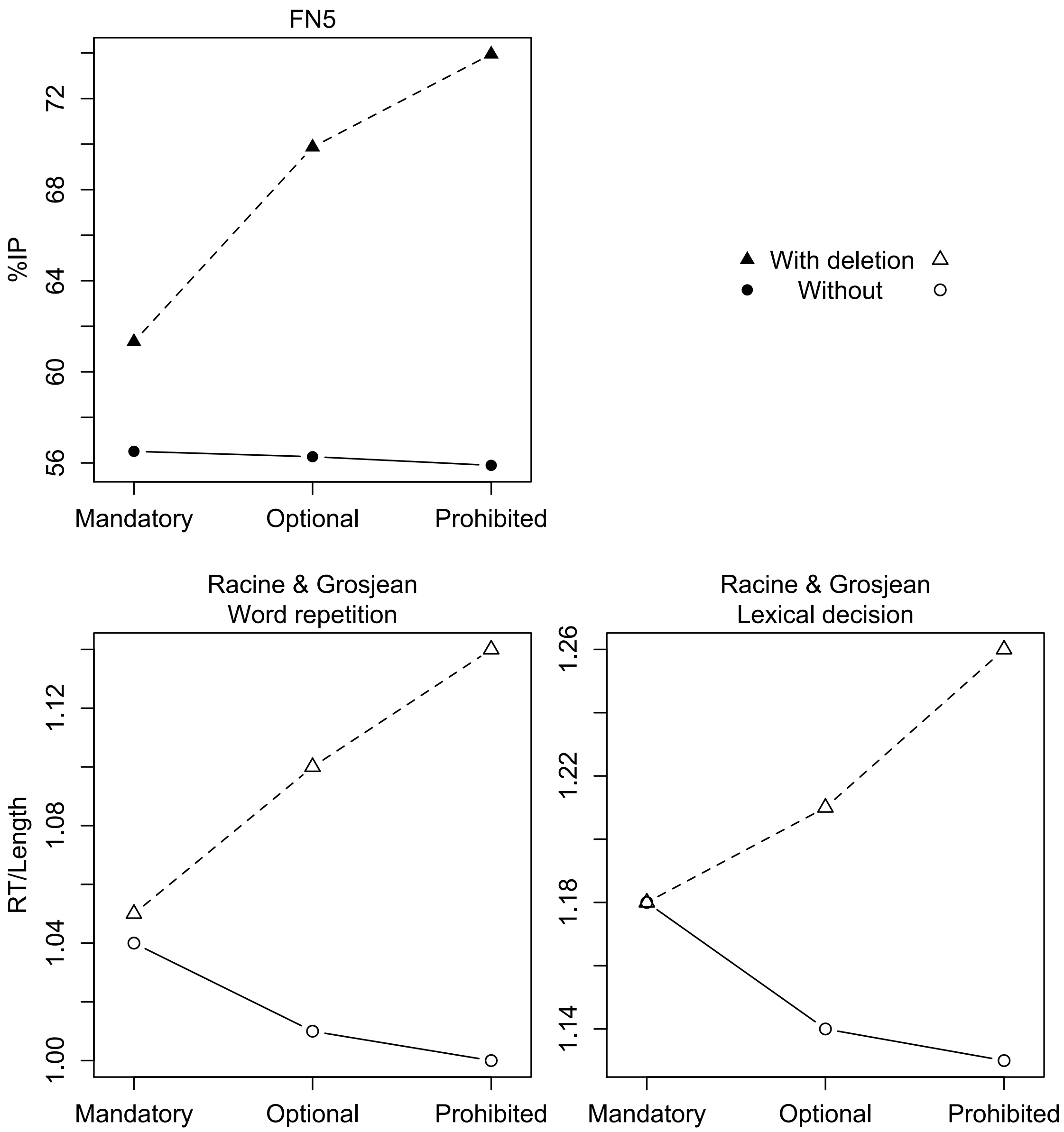


Figure 19. Comparing the results of FN5 with the results of the two perception experiments by Racine and Grosjean (2005).

linking with liaison, the ending consonant is latent and only pronounced when it precedes a word beginning with a vowel or glide (e.g. "mon~ air"). (As clarified in fn. 13, there are the exceptions of aspirated *h* and certain words with glides.) In linking without liaison, the ending consonant is always pronounced, both before a vowel or glide (e.g. "même arbre") and before another consonant (e.g. "chaque rue"). Yersin-Besson and Grosjean used 24 sequences of two words (always an adjective or a determiner followed by a noun) and had 8 sequences for each of the three types of linking: with liaison C to V, without liaison C to V, and without liaison C to C. To each sequence with linking, there corresponded a distractor sequence without linking: such as, “mon~ air” - “mon nerf”, “même arbre” - “même marbre”, and “chaque rue” - “chaque crue”. According to Yersin-Besson and Grosjean’s results, a sequence like “mon~ air” (linking with liaison) is recognized with considerable difficulty and delay, because it is phonemically totally ambiguous to its counterpart without linking, “mon nerf”. (Where should the consonant be attached: to the end of the first word or to the beginning of the second word?). By contrast, sequences like “même arbre” and “chaque rue” (linking without liaison) are recognized without any difficulty or delay, since they are not phonemically ambiguous to their counterparts without linking, “même marbre” and “chaque crue”. (There is one consonant for the former sequences but a double consonant for the latter sequences.)

To test FN5 on linking stimuli of the kind of Yersin-Besson and Grosjean (1996), we had to load a special lexicon file (“F.PlurAnts.Lex”), which contains 34 numerals and other adjective forms only used before plural nouns. In fact, Yersin-Besson and Grosjean’s stimuli included some in the plural (“mille armes”, “leurs~ ailes”, etc.). Since FN5 neither has the nouns in the plural, we converted those concerned to singular nouns; they are pronounced identically (e.g. “arme” and “aile”). We verified whether Yersin-Besson and Grosjean’s 3 × 8 sequences have the intended characteristics in FN5’s lexicon, among other things, that the 8 sequences with liaison are phonemically ambiguous (this was the case) and that the 16 sequences without liaison are not, but we found that 7 sequences without liaison are phonemically ambiguous as well (e.g. “fausse heure” = *“faux sœur”, “grande anse” = *“grand danse”, “longue rêne” = *“long graine”). They are not ambiguous for humans, but for the model, because it does not account for gender agreement. To correct this problem, we set up a modified list of 3 × 8 stimuli, in which we substituted the offending adjective or determiner by another one with the same linking consonant (e.g. “dense heure” with /s/ like “fausse”, “fade anse” with /d/ like “grande”, “vague rêne” with /g/ like “longue”). Finally, as 3 × 8 stimuli is a small sampling size (we will see that it did not quite suffice to produce a significant effect), we also prepared a larger list. It includes all stimuli of the modified list but extends it to 3 × 20 stimuli (i.e. 20 sequences with liaison, 20 sequences without liaison C to V, and 20 sequences without liaison C to C). All the word properties (frequency, length, uniqueness point, etc.) were strictly controlled.

Using these lists of stimuli (original, modified, and extended version), we performed three times the same evaluation. We first ran the sequences (on the full standard French lexicon, plus “F.PlurAnts.Lex”) and established, as usually, the number of sequences isolated as well as the nouns’ %IP in each sequence.

We then reran the nouns alone, this time on a lexicon of only nouns (“F.Nouns-YBG96.Lex”, excluding nouns with aspirated *h*), and determined the nouns’ %IP in isolation. This allowed us to compute the difference of the %IP in sequence minus the %IP in isolation; it expresses by how much a noun is held back when the noun is perceived in that particular sequence compared to when it is heard in isolation. When the %IP difference is 0, the noun takes the same time in the sequence as in isolation; if the difference is greater than 0, the noun is delayed in the sequence; and if the difference is less than 0, the noun is advanced in the sequence (the last situation is rare but not impossible). Running the original list of 3 × 8 sequences of Yersin-Besson and Grosjean, we found that 6 sequences (4 with liaison, 1 without liaison C to V, and 1 without liaison C to C) were not isolated by the model: it activated “ton nombre” for “ton~ ombre”, “son neuf” for “son~ œuf”, “cent tasse(s)” for “cent~ as”, “leur zèle” for “leurs~ aile(s)”, as well as *“faux sœur” for “fausse heure”, and *“long graine” for “longue rêne”. They were omitted from the analysis. For all the other sequences, the %IP difference (%IP in sequence − %IP in isolation) was calculated. It was on average 3.776% (*SD* = 7.358%) for the sequences with liaison, 5.923% (*SD* = 13.636%) for the sequences without liaison C to V, and 0% all the way through for the sequences without liaison C to C. There was no effect of type of linking: $F(2, 15) = 0.733$, *MSE* = 85.202, *NS*. The simulation results agree only partially with the results of Yersin-Besson and Grosjean. We did find that the sequences with liaison (e.g. “ton~ ombre”) are difficult to recognize: they are phonemically identical to their counterpart without linking (“ton nombre”); half of them were isolated correctly (i.e. with linking), even if delayed (%IP difference > 0), and half of them were not isolated (their counterpart without linking was instead proposed). Regarding the sequences without liaison, we mentioned that 7 of them are phonemically ambiguous for the model but not for the human, since FN5 does not account for gender agreement. It is of no surprise that those sequences (e.g. “fausse heure” and “grande anse”) are influenced by distractor words (like “sœur” or “danse”), and thus either momentarily delayed or totally impeded, in much the same way as the sequences with liaison.

We now ran the modified list of 3 × 8 sequences, in which, phonemically, all sequences with liaison are ambiguous but all sequences without liaison are non-ambiguous, and had the following results. Again, half of the sequences with liaison were not isolated by the model (they are the same 4 sequences as in the original list: “ton~ ombre”, “son~ œuf”, “cent~ as”, “leurs~ aile(s)”); they were activated as their counterpart without linking (“ton nombre”, etc.), and removed from further analysis. All 16 sequences without liaison were now isolated by the model (unlike the original list, those in the modified list are all non-ambiguous). The %IP difference (%IP in sequence − %IP in isolation) was, on the average, 3.776% (*SD* = 7.358%) for the sequences with liaison (same items as above), −1.654% (*SD* = 2.066%) for the sequences without liaison C to V (3 new items), and again all 0% for the sequences without liaison C to C (4 new items). There was a marginally non-significant effect of type of linking: $F(2, 17) = 3.480$, $MSE = 11.311$, $p = .054$. Since this evaluation concerned a small sample size (3 × 8), we had reasonable hope that we would find a significant effect using a larger sample size. So we finally ran our extended list of 3 × 20 sequences. Also here, half of the sequences with liaison (i.e. 10 sequences) were not isolated by FN5 (besides the 4 above, they are: “léger~ ail”, “grand~ an”, “long~ arçon”, “cent~ ire(s)”, “son~ once”, and “cent~ urne(s)”), but all the sequences without liaison (i.e. 40 sequences) were isolated by the model. The mean %IP difference was 5.281% (*SD* = 8.820%) for the sequences with liaison, −0.141% (*SD* = 2.382%) for the sequences without liaison C to V, and 0.625% (*SD* = 2.841%) for the sequences without liaison C to C. And indeed, there was a significant effect of type of linking: $F(2, 47) = 5.110$, $MSE = 20.452$, $p < .01$. Post-hoc analyses revealed a difference between the sequences with liaison and the sequences without liaison ($p < .01$ for the type C to V, $p < .05$ for the type C to C) but not between the two types of sequences without liaison. We arrive at the conclusion that the simulation results now agree well with the results of Yersin-Besson and Grosjean. Sequences with liaison are recognized with much more difficulty than sequences without liaison: half of the former, but all of the latter, are isolated by the model. Words after a linking with liaison are recognized more slowly than words after a linking without liaison: the nouns in the sequences with liaison are isolated with a sizable delay (of about 5% of their word length) as compared to the nouns in the sequences without liaison (no delay).

Two final examples, “tout~ est” (one of our sequences with liaison) and “chaque ours” (one without liaison), are shown in Figure 20. In the upper panel, the adjective “tout” (in its pronunciation variant /tut/, i.e. the liaison form) and the noun “est” are activated. We find that “est” is isolated very late, and barely so. Although “est” is proposed at the very moment shown (i.e. at the end of silence), we observe that “test” is nearly as strong as “est”. In fact, would we add just a few additional cycles, “test” would surpass “est”, and “tout” would then take its pronunciation variant /tu/ (non-liaison form). Of course, “tout~ est” (with linking) and “tout test” (without linking) are phonemically totally ambiguous, and hence it is only natural that FN5 has a hard time deciding where to put the consonant /t/ (to the end of “tout~” or to the start of “test”) and choosing between one and the

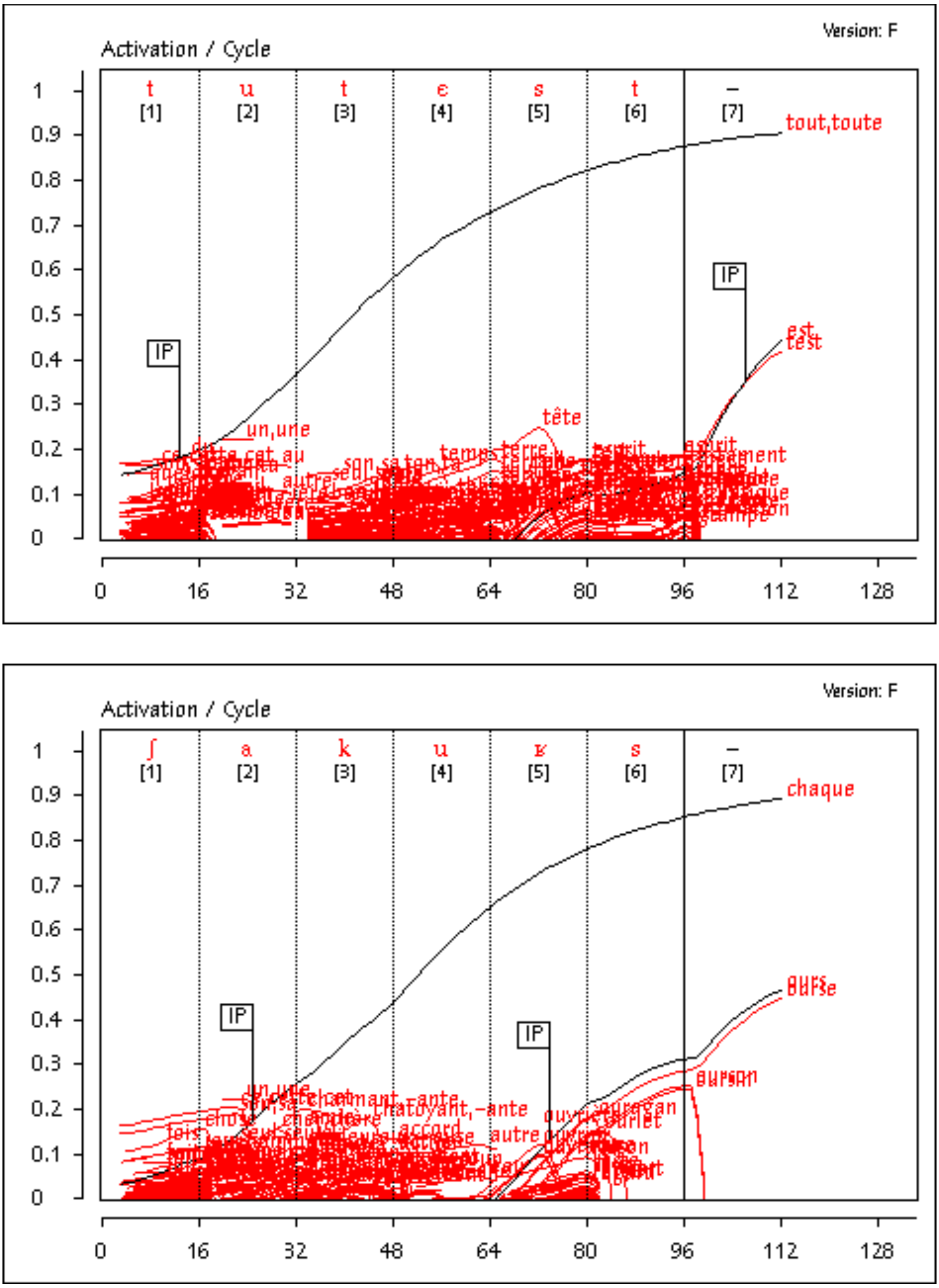


Figure 20. Simulating the recognition of “tout~ est” (/tutɛst/), two words linked with liaison, and of “chaque ours” (/ʃakuʁs/), two words linked without liaison.

other valid segmentation of /tutɛst/ (which one the model settles on depends on several factors, including the words' frequency). By contrast, "chaque ours", in the lower panel, faces no such difficulties: it is a non-ambiguous sequence, and hence the words "chaque" and "ours" are proposed straightaway. (The potential distractor sequence "chaque course" is never once considered because it would require a /kk/ instead of the one /k/ of "chaque ours".) Now let us examine the timings in the two examples. For the word "est" in the upper example, the %IP in the sequence "tout~ est" is 90.63%, whereas the %IP in isolation is 78.13%; the %IP difference (in sequence minus in isolation) is therefore 12.50%, which means that "est" is indeed delayed in this sequence with liaison. By contrast, for the word "ours" in the lower example, the %IP in the sequence "chaque ours" is 40.63%, and the %IP in isolation is exactly the same; the %IP difference of 0% signifies that "ours" is not at all delayed in this sequence without liaison.

CHAPTER 7. EVALUATING BIMOLA

We now move on to the evaluations of BIMOLA, our model of bilingual spoken word recognition. They follow the methodology that we have introduced in the evaluations of FN5 (at the beginning of the previous chapter). That is to say, also for BIMOLA, there are large-scale tests that examine the model's general (i.e. overall) recognition performance, and there are other, smaller tests that focus on just a few but carefully selected words: parametric studies that serve to find out whether the model can account for specific psycholinguistic effects. As shown in Table 10, we will report both monolingual and bilingual simulations (all on single words since BIMOLA, unlike FN5, does not process sequences of words), and in either section, we will describe first the general recognition and then the specific effects. As we did throughout the evaluations of FN5, we will again, from time to time, visualize typical examples of word recognitions in BIMOLA graphically but, apart from that, we will of course mostly present the results and interpretations of statistical analyses. All the evaluations were run using the default setting of parameters ("Bimola.Diss.Set"). Once more, our macro functionality came in handy to present all the many words to the model. To allow other researchers to replicate the BIMOLA evaluations or part of them, a list of the macros used can be found in Appendix C2.

Table 10.
Outline of the BIMOLA evaluations

Monolingual simulations
General monolingual recognition Comparison with FN5 Monolingual effects: - Frequency - Length - Uniqueness point
Bilingual simulations
General bilingual recognition Effects in bilingual guest word recognition: - Language phonetics - Phonotactics - Near-homophony

Monolingual simulations

Even though BIMOLA is primarily a bilingual, not a monolingual, model of spoken word recognition, it can very well function with only one language, English or French. In our monolingual simulations, we will first look into monolingual BIMOLA's overall ability to recognize words, be it in English or in French. We will then draw a comparison between monolingual BIMOLA and our first model, the monolingual model FN5. Lastly, we will briefly revisit three specific psycholinguistic effects (frequency, length, and uniqueness point) that we already investigated in the evaluation of FN5. Indeed, it is important to ascertain that BIMOLA, too, can simulate monolingual effects.

General monolingual recognition. Each and every word of BIMOLA was tested in this evaluation.[40] First, the English lexicon was loaded and BIMOLA

[40] We could do this because BIMOLA deals with single words only, which all have the one same alignment position, and therefore runs very fast. Compare this to FN5, which processes multiple words presented sequentially, and for which we selected 1,000 of its 17,668 words for the general recognition evaluation.

was run as an English-only monolingual model, operating exclusively with English phoneme units and English word units. One after another, BIMOLA's 4,348 English words (base-form verbs) were presented to the model, always with a silence input after the end of the word, but without any additional cycles. Subsequently, the English lexicon was replaced with the French lexicon and BIMOLA was run as a French-only monolingual model, consisting instead of French phoneme units and French word units. BIMOLA's 4,348 French words (they are verbs in the 3rd-person singular subjunctive) were now presented to the model, again including a final silence input but no additional cycles. It was found that absolutely all the 8,696 words were correctly isolated by BIMOLA; that is, the general recognition success rate was a perfect 100%. This should come as no surprise. In BIMOLA, all words have the same alignment position (there is no segmentation into multiple words) and, as homophones are absent in both of BIMOLA's monolingual lexicons, all words are sure to become unique by the end of the silence input (even words embedded in longer words).

Comparison with FN5. It is useful to compare French-only monolingual BIMOLA with our other model, FN5, which we specially built to simulate French monolingual spoken word recognition. Any differences in behavior between the two models could either be caused by certain model-specific assumptions (e.g. in the architecture, internal mechanisms, parameter settings, etc.) or they could simply stem from the respective French lexicon used by the model (a lexicon of 4,348 verbs in BIMOLA, but one of 17,668 nouns, determiners, and prenominal adjectives in FN5, the two containing quite different linguistic information).[41] To find this out, we tested the 901 French words that overlap between BIMOLA's and FN5's lexicons (these words are 3rd sing. subjunctive verbs in BIMOLA and they are nouns or adjectives in FN5, such as the verb "agence", from "agencer", vs. the noun "(une) agence"). We first ran these words two times in monolingual BIMOLA, on one occasion we had loaded BIMOLA's French lexicon and on the other occasion we had loaded FN5's French lexicon (standard French version). We then ran the identical list of test words twice in FN5, also here once having loaded BIMOLA's and once having loaded FN5's French lexicon.[42]

---

[41] Michael Thomas (personal communication, 2014) suggested us this comparison.

[42] Note that in the case of multiple pronunciation variants, BIMOLA uses only the first variant, disregards the rest, and thus cannot recognize words with deleted schwa, nor feminine forms of adjectives; 23 items were affected. FN5 recognizes all but 3 very short items, which would need a few more cycles to reach the isolation point.

Our test words can be plotted as points showing the result in BIMOLA on the y-axis and the result in FN5 on the x-axis. This was carried out, a total of four times, for Figure 21. The panel on the top left shows the relationship when each monolingual model operates on its own lexicon (BIMOLA on the BIMOLA lexicon, FN5 on the FN5 lexicon). As one can see in this panel, there is a vast dissimilarity between the results of BIMOLA and FN5; the Pearson correlation coefficient r is 0.328 (all p's < .001); the linear regression line (the dotted line) does not lie on the diagonal but is biased downward. For only 256 of the 901 test words, the result differs by less than 5%. The words that display the most

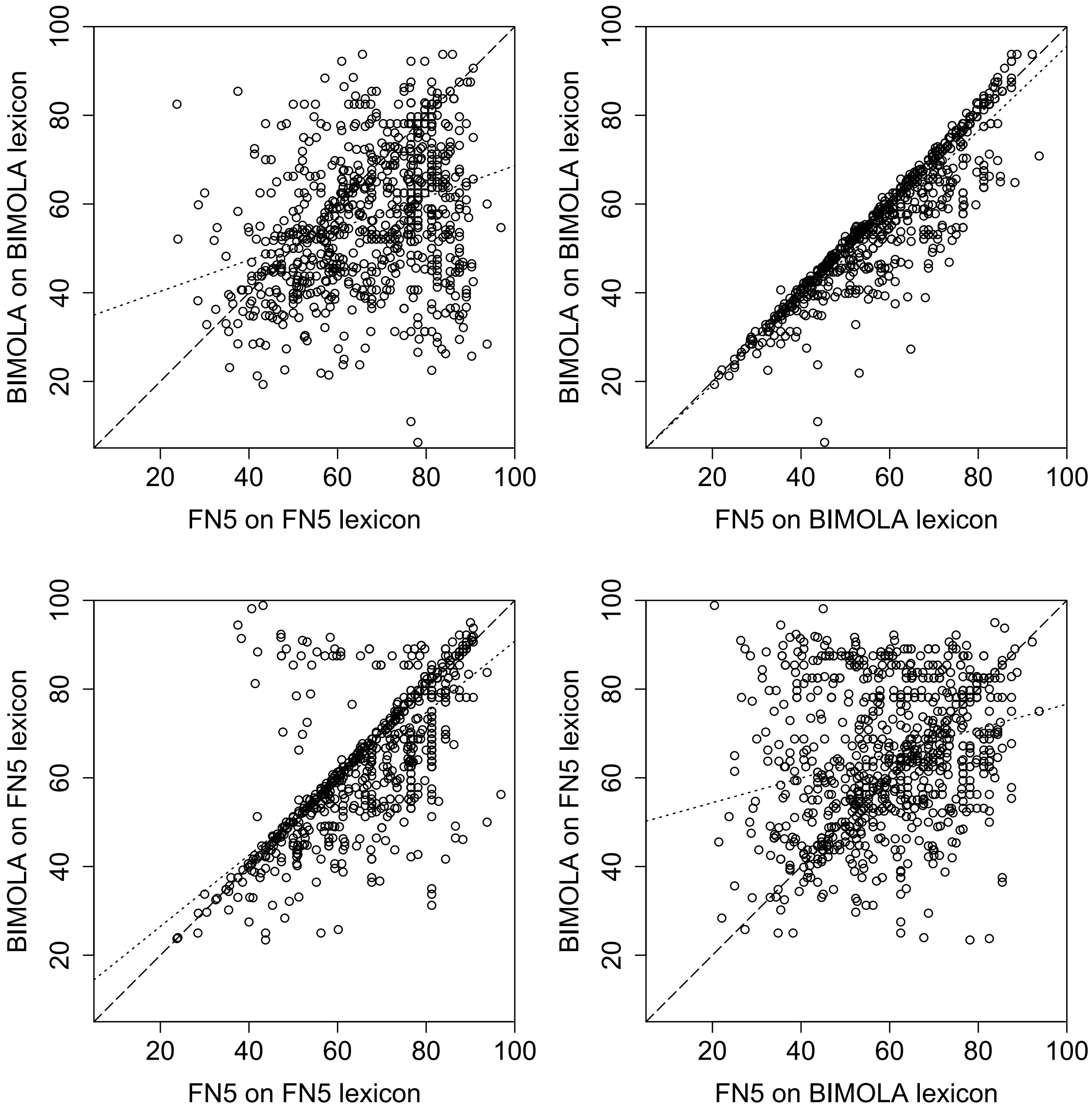


Figure 21. Relationship of results, on the same French words, between FN5 and the French-only monolingual BIMOLA, depending on whether the models operate on FN5's or BIMOLA's French lexicon. The measurement is the %IP.

dissimilar results are typically those that have the most dissimilar characteristics in BIMOLA's versus FN5's lexicon. For example, "gouverne" is a medium-high frequency verb with an early uniqueness point in BIMOLA's lexicon, but is a low frequency noun with a late uniqueness point in FN5's lexicon; no wonder that it is one of the words that are isolated considerably faster by BIMOLA than FN5. In contrast, "affaire" is a low frequency verb in BIMOLA's lexicon (embedded in and a little less frequent than "affermisse"), but it is a very high frequency noun in FN5's lexicon (more frequent than any word that starts with /af/); and so, this word is isolated a great deal more slowly by BIMOLA than by FN5.

All the differences between BIMOLA and FN5 that have their origin in the lexicons are eliminated in the next two panels. When both models operate on the BIMOLA lexicon (as visualized in the top right panel), the results of BIMOLA and FN5 become very similar; now, the correlation coefficient *r* is 0.922, and the regression line nearly coincides with the diagonal. For as many as 710 of the 901 words, the result differs by less than 5% (including "gouverne" and "affaire", which now are isolated exactly as quickly by BIMOLA and FN5). Still, for some words, the result differs by more than 5%; they are predominantly located below the diagonal, meaning that they are isolated more quickly by BIMOLA than FN5. We identified three main causes. First, in FN5, a top-frequency word ("ait", "soit", "dise", etc.) may briefly turn up at a later alignment position and delay the target word's isolation; in BIMOLA, all words (also top-frequency words) are aligned at the start and never reappear at later positions. Second, words with a short word embedded at onset ("niche" - "nie", "voyage" - "voie") may be isolated about one phoneme later in FN5 (owing to words without continuation inhibition, see p. 79) than in BIMOLA (silence phoneme inhibition, cf. the footnote on the same page). Third, words with a competitor of equal frequency ("cravate" - "cravache") may surpass the competitor just a little and are then said to be isolated in BIMOLA but not yet so in FN5 (parameter "Difference for isolation point": 0 in BIMOLA, 0.001 in FN5). The bottom left panel presents the analogous (if less clear-cut) situation when both models operate on the FN5 lexicon; here, the correlation coefficient *r* is 0.714, and the regression line is approximately on the diagonal. For 549 of the 901 words, the result differs by less than 5% (0% for "gouverne" and "affaire"). There are again the outliers below the diagonal, for the reasons mentioned before (here with top-frequency determiners). There are also a few outliers above the diagonal (words isolated more slowly by BIMOLA than FN5). In BIMOLA, top-down feedback from words to phonemes and up to words may activate by mistake a very long competitor word and delay the target's isolation (e.g. "orchestre" - "orchestration", "anesthésie" - "anesthésiste/anesthésique").

To sum up, this model comparison was instructive in that it revealed that French-only monolingual BIMOLA and FN5 behave very similarly and produce very similar results, as long as they are tested on the same monolingual lexicon. The result differences we found were attributed to model-specific assumptions, regarding the architecture (top-down activation: on in BIMOLA, off in FN5), the internal mechanisms (position processor: absent in BIMOLA, present in FN5; influence of following context: simple mechanism in BIMOLA, more general one in FN5), and a parameter setting (isolation point: zero threshold in BIMOLA, small threshold in FN5). But French-only monolingual BIMOLA and FN5 behave very differently when each uses its own lexicon (as they normally do, of course). This also true, as shown in the bottom right panel, when each model operates on the other one's lexicon (BIMOLA on the FN5 lexicon, FN5 on the BIMOLA lexicon). Once again, as a consequence of the disparity between the lexicons, there is huge variation between the results of the models (r = 0.259; regression line slants; result differs by less than 5% for only 295 of the 901 words).

We will focus on English words, for a change (not French words, on which we already spent pretty much time with FN5), and thus we will work with the English-only BIMOLA for the remaining monolingual simulations.

Three effects revisited. Frequency, length, and uniqueness point (UP) are three of the most fundamental phenomena in spoken word recognition (see the literature we cited when we studied these language-independent effects in FN5, one effect at a time; pp. 98ff). In BIMOLA, we examined the three effects all together by choosing 64 English test words with the following properties: Half of the words are of low frequency (value in the lexicon < 0.17, M = 0.097, SD = 0.063) and half are of high frequency (value in the lexicon > 0.35, M = 0.454, SD = 0.067); half are short words (composed of three phonemes) and half are long words (consisting of six phonemes); half have an early UP (at 66% of the word, i.e. on the second of three or on the fourth of six phonemes, depending on the word's length) and half have a late UP (at 100% of the word, i.e. always on the last phoneme). By using eight words in each cell of the 2 × 2 × 2 design, the three variables were mutually controlled.

We ran these stimuli in English-only BIMOLA, and we determined their IP@cycle (isolation point in absolute simulation cycles) and %IP (isolation point in percentage of the word length including the final silence input; see p. 66f). Each measure was entered into a three-way analysis of variance. As concerns the measure of IP@cycle, main effects were found for the factor of frequency, $F(1, 56) = 21.869$, for the factor of length, $F(1, 56) = 112.557$, and for the factor of UP, $F(1, 56) = 40.126$, $MSE = 93.114$, all three p's < .001. There were no interactions between the factors (all F's < 2.0). The cell means (and *SD*'s) are: low/short/early: 30.5 (4.140), high/short/early: 18.25 (10.039), low/long/early: 58.5 (4.536), high/long/early: 41.5 (7.801), low/short/late: 43.875 (6.266), high/short/late: 35.5 (7.483), low/long/late: 69 (14.122), high/long/late: 61.5 (15.838). The measure of %IP neutralizes the length of a word; so it is not surprising that it showed main effects for the factor of frequency, $F(1, 56) = 24.795$, as well as for the factor of UP, $F(1, 56) = 47.757$, $MSE = 118.047$, both p's < .001, but not for the factor of length, $F(1, 56) = 0.266$, *NS*. Again there were no interactions between factors (all F's < 3.7). The cell means (and *SD*'s) are: low/short/early: 47.656% (6.469%), high/short/early: 28.516% (15.686%), low/long/early: 52.232% (4.050%), high/long/early: 37.054% (6.965%), low/short/late: 68.555% (9.791%), high/short/late: 55.469% (11.693%), low/long/late: 61.607% (12.609%), high/long/late: 54.911% (14.141%). We conclude that these three monolingual effects are simulated by BIMOLA.

Figure 22 visualizes eight examples of word recognitions in the model, one word for each of the eight cells of the Frequency × Length × UP evaluation. In each cell, we selected the word whose isolation point (IP@cycle) is closest to the mean: “ogle” with 28, “offer” with 17, “reaffirm” with 59, “direct” with 41, “lout” with 44, “pull” with 36, “inflate” with 77, and “destroy” with 60. Just as in the FN5 diagrams, activation levels of the candidates proposed by BIMOLA are shown in function of simulation cycles; the target words (“ogle”, “offer”, etc.) are drawn in black[43] (with a flag indicating the IP) and the other English candidates in blue. The three effects under discussion can all be easily observed by contrasting the diagrams in pairs. For the effect of frequency (low-frequency words are isolated more slowly than high-frequency words), we compare first-line with second-line panels and third-line with fourth-line panels: “ogle” vs. “offer”, “lout” vs. “pull”, “reaffirm” vs. “direct”, and “inflate” vs. “destroy”. To see the effect of UP (words with early UP are isolated more quickly than words with late UP), we compare panels on the left with panels on the right: “ogle” - “lout”, “offer” - “pull”, “reaffirm” - “inflate”, and “direct” - “destroy”. Finally, to observe the effect of length (short words are isolated faster than long words), we compare first-line with third-line panels as well as second-line with fourth-line panels: “ogle” - “reaffirm”, “offer” - “direct”, “lout” - “inflate”, and “pull” - “destroy”. (Since the x-axis is adjusted to word length, the effect of length may seem to disappear graphically; this is not the case, but corresponds to what is meant by the measure of %IP.)

---

[43] In white (on a gray background) in the simulation program.

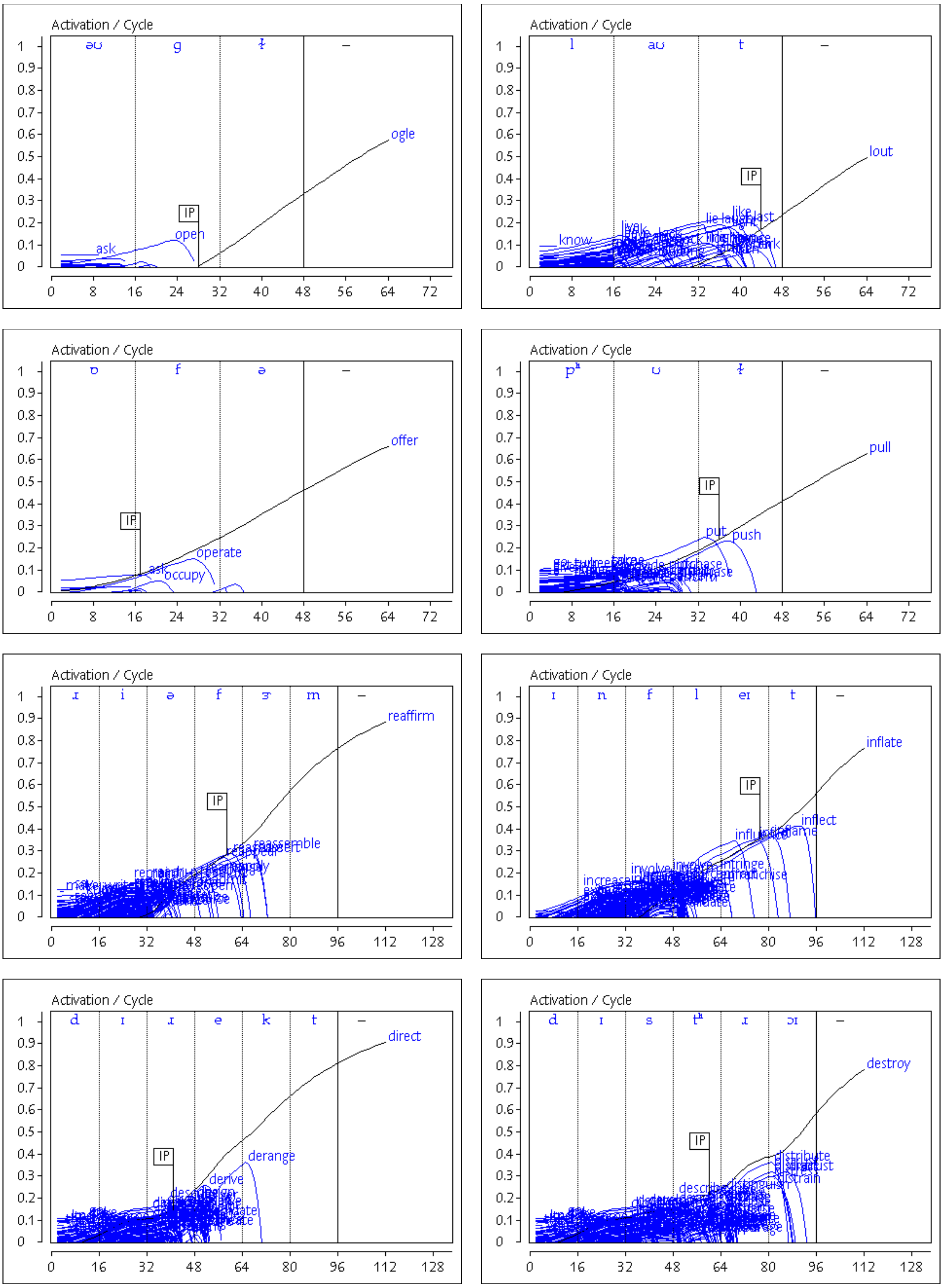


Figure 22. Running eight words in the English-only monolingual BIMOLA: “ogle”, “offer”, “reaffirm”, “direct”, “lout”, “pull”, “inflate”, and “destroy”.

Bilingual simulations

Of course, BIMOLA normally operates as a bilingual model of spoken word recognition. It then deals simultaneously with two languages, English and French, of which always one serves as the base language and the other as the guest language. In our bilingual simulations, we will first establish that bilingual BIMOLA can recognize words from either language, and we will contrast base language words with guest words (general bilingual recognition). We will show subsequently that BIMOLA simulates a number of effects that have to do with the specific properties of guest words (language phonetics, phonotactics, and near-homophony).

General bilingual recognition. Can BIMOLA function with either French or English as the base language, and is it able to identify words from the base language as well as from the guest language? To put these claims to the test, the bilingual lexicon, which combines both the English and the French lexicons, was loaded and BIMOLA was run as a bilingual model, with its two subsets of phoneme units and two subsets of word units. In the first part of the evaluation, French was the base language (at a language activation level of 100%) and English was the guest language (at a language activation level of 80%, which represents a bilingual language mode). All of BIMOLA's 4,348 French words were tested as base language words, and all of BIMOLA's 4,348 English words were tested as guest words, that is, as code-switches from English into French. The whole situation was symmetrically reversed for the evaluation's second part. Now, English served as the base language (at 100% language activation level), while French served as the guest language (at 80% language activation level). Thus, the same 4,348 English words were here tested as base language words, whereas the same 4,348 French words were here tested as guest words (code-switches from French into English). Every time we evaluated a word in bilingual BIMOLA, we used not only the final silence input but also ran 16 extra cycles at the very end of the simulation. It counted as a recognition success if the correct word exceeded the response strength of all other words (in the same language and in the other language) and if it remained at the top until the very end of the simulation. But when a word only achieved its top place, or lost it, during the 16 extra cycles, or never gained it at all, it counted as a recognition failure.

We had the following results when French acted as the base language. The recognition success rate (= successful items ⁄ number of items × 100%) was 99.95% for base language words (French) and 97.06% for guest words

(English), meaning that only 2 of all the 4,348 French words and 128 out of all the 4,348 English words were not isolated by BIMOLA. We examined each of the words that failed and found that the two base language words, French “mine” and “scie”, which are of very low frequency (value in the lexicon = 0.068), were confounded with English “mean” and “see”, which are of very high frequency (value in the lexicon = 0.561 for “mean” and 0.692 for “see”). French “mine” and English “mean”, which were activated in BIMOLA one along with the other, are pronounced very much alike (the two are transcribed /min/ and they have, in our bilingual metric space of phonemes, a distance sum of 0 + 1.64 + 0.10 = 1.74). Likewise, French “scie” and English “see”, which again were both activated in BIMOLA, are pronounced very similarly (transcription /si/, distance sum of 1.64). These pairs of words are near-homophones across the two languages, which a bilingual model should not (and cannot) decide on without having recourse to semantic context. Cross-language near-homophones were also a main reason why BIMOLA did not isolate all guest words. For 72 guest words (English code-switches), a French near-homophone was activated as well: some kept the top place even during the additional cycles at the end (e.g. F “cloue” for E “clew”, F “moque” for E “mock”, F “coiffe” for E “quaff”, etc.), a few finally backed down (e.g. F “passe” for E “patch” as this pair is indeed quite dissimilar), but others took newly over (e.g. F “laque” for E “lack”). Another common source of error (for 50 guest words) were high-frequency distractors that were more activated than the low-frequency targets (e.g. “allow” for “ally”, “close” for “clothe”, “get” for “gape”, “love” for “luff”, “tell” for “till”, etc.); the low-frequency targets all caught up during the extra cycles at the end. Finally, for 3 words, there was a combination of both reasons, and the 3 one-phoneme words “awe”, “eye”, and “owe” simply did not get enough time to become active (again, extra cycles help).

The results were analogous when English served as the base language. The recognition success rate was 99.95% for base language words (English) and 98.28% for guest words (French). That is, just 2 out of all the 4,348 English words and 75 of all the 4,348 French words were not isolated by BIMOLA. The base language words that did not make it (again two, as chance had it) were “dish” and “swear”. The former was mixed up with French “dise” and the latter was mistaken for French “soit”, understandably so because “dise” and “soit” are among the very most frequent verbs of French (values in the lexicon = 0.812 for “dise” and 1 for “soit”). But since these pairs of words are not pronounced the same (distance sum of 1.1 + 2.6 + 1.8 = 5.5 for “dish” - “dise”, and 0 + 0 + 2.6 = 2.6 for “swear” - “soit”), they can easily be set apart with a few additional cycles.

As regards the guest words that were not isolated by BIMOLA, cross-language near-homophones again played an important role. For 38 guest words (French code-switches), an English near-homophone was also activated. We met again 15 words (such as E “nap” for F “nappe”, E “mash” for F “mâche”, and E “shock” for F “choque”) that we had already encountered in the reversed condition when French near-homophones were co-activated for English guest words (F “nappe” for E “nap”, F “mâche” for E “mash”, and F “choque” for E “shock”). But several that had appeared there (such as F “cloue” for E “clew”) did not reoccur here, and instead some were new (E “boom” was activated for F “boume”, E “meet” was activated for F “mite”, and so on). The reason for the asymmetry has to do with the frequency pull between words (Grosjean, 1988), that is, if it is the base language word or the guest word that is the more frequent one (e.g. F “cloue” is more frequent than E “clew”, but F “mite” is less frequent than E “meet”), or whether both words of the pair have about the same frequency (as is the case for F “nappe”, “mâche”, “choque” vs. E “nap”, “mash”, “shock”). Besides cross-language near-homophones, failure types observed for English guest words in French were found again for French guest words in English: single-phoneme items that did not have sufficient time to get going (1 word, F “hue”), and high-frequency distractors that were more activated than the low-frequency targets (36 words concerned; e.g. “baisse” for “bêche”, “dise” for “bise”, “déserte” for “désherbe”, “gêne” for “geigne”, “juge” for “juche”, “tende” for “tangue”, etc.; most of them succeeded to catch up during the 16 additional cycles; some would need more).

BIMOLA’s general recognition success rates are recapitulated in the left half of Table 11. They are organized here not by which is the base language and which is the guest language (i.e. in the way we have described so far) but are rearranged according to the items tested (English words vs. French words). In truth, the words tested were exactly the same ones in each condition. As a consequence, the results of BIMOLA as a bilingual model can be juxtaposed with those of BIMOLA as a monolingual model, which we reported earlier on. Medians, means, and standard deviations of the %IP are provided in the right half of the table. As one can see, when words were tested as base language words in bilingual BIMOLA, they were isolated just as rapidly as when they were tested in monolingual BIMOLA: the median %IP is identical (54.69% for English words, 53.12% for French words) and the mean %IP differs barely, by 0.45% of word length for English words and by 0.25% of word length for French words (these differences, albeit tiny, are significant due to the huge sample sizes: paired $t(4345) = 12.47$ and 10.76, both p’s $< .001$). However, when words were run as guest words in bilingual BIMOLA, they were isolated more slowly: the median %IP is higher (62.50% for English words, 58.59% for French words) and the mean %IP differs noticeably from the one in monolingual BIMOLA, by 6.65% of word length for English words and by 5.46% of word length for French words (paired $t(4219) = 60.27$ and paired $t(4272) = 56.88$, both p's < .001).

In sum, bilingual BIMOLA is indeed able to identify base language words and guest words, regardless of which is the base language, English or French. Base language words are recognized as fast as the same words in monolingual BIMOLA, but guest words take usually a little longer (and in fact, bilinguals need more time to access guest words than base language words; see the results of Soares & Grosjean, 1984).

Table 11.
General recognition success rates and isolation points of BIMOLA

	Success rates		*Isolation points (%IP)*	
	English words	*French words*	*English words*	*French words*
BIMOLA as monolingual model	100%	100%	*Mdn* 54.69% *M* 54.31% *SD* 15.37%	*Mdn* 53.12% *M* 53.88% *SD* 14.84%
BIMOLA as bilingual model				
Words run as base language words	99.95%	99.95%	*Mdn* 54.69% *M* 54.76% *SD* 15.42%	*Mdn* 53.12% *M* 54.13% *SD* 14.92%
Words run as guest words	97.06%	98.28%	*Mdn* 62.50% *M* 60.97% *SD* 15.20%	*Mdn* 58.59% *M* 59.34% *SD* 14.16%

Let us demonstrate concretely two examples of words, one English and one French, in Figure 23. The two words, English “sing” and French “tâche”, have a similar frequency and the same length and uniqueness point. (The “H” in the label of “tâche(H)” reminds us that it also stands for homophonous “tache”.) Both words were first run in the appropriate monolingual BIMOLA (English-only for the English word, French-only for the French word) and were then run in a

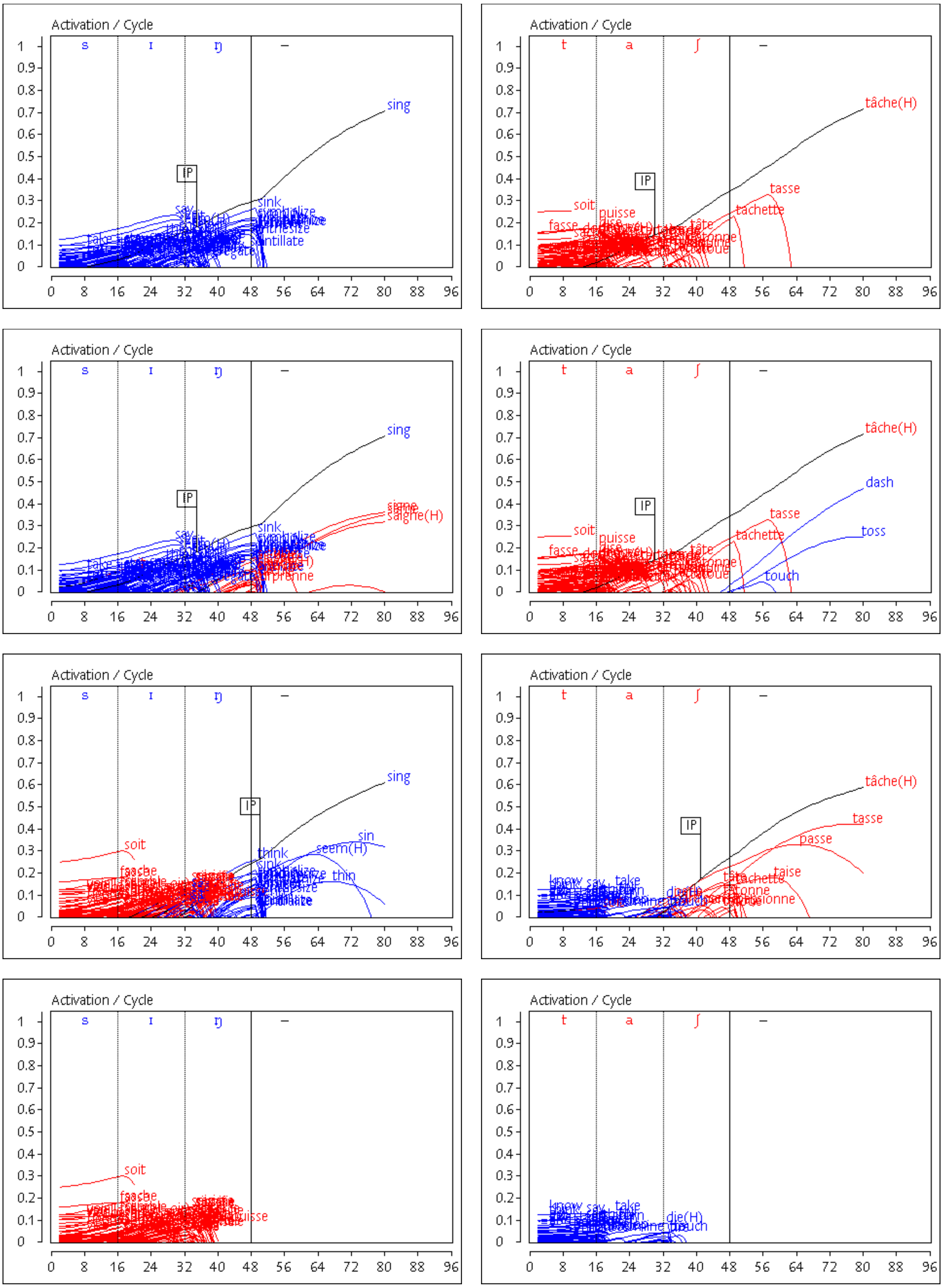


Figure 23. Running English "sing" and French "tâche" (/taʃ/) in the appropriate monolingual BIMOLA (top panels), either as base language words (second-line panels) or as guest words (third-line panels) in a bilingual BIMOLA, and lastly in the inappropriate monolingual BIMOLA (bottom panels).

bilingual BIMOLA, once as base language words and once as guest words. In monolingual BIMOLA (see the two panels at the top of the figure), candidates are activated obviously in only one language, the target word among them (E “sing” and F “tâche”, respectively). English is drawn in blue, French in red. As for base language word recognition (the second row of panels), we observe the following: the initial candidates are always from the base language (i.e. the same language as in the top row); the target word is one of these candidates and is isolated at exactly the same point as in monolingual BIMOLA (IP@cycle = 35 for E “sing” and 30 for F “tâche”); candidates from the guest language are activated as well (including F “signe”, “saigne”, “sème” for E “sing”, and E “dash, “toss”, “touch” for F “tâche”) but later and they do not play an important role. Also in guest word recognition (the third row of panels), the candidates first proposed are from the base language (e.g. F “soit” or “sache” for E “sing”, and E “take” or “think” for F “tâche”), but since it is not the same language as in the top row, the target word is not among them; after some time, guest language candidates appear as well (they begin from a lower resting level),[44] and thus the target word is isolated at a later moment (IP@cycle = 50 for E “sing” and 41 for F “tâche”). To complete the picture, the two words were finally tested in what is for them the inappropriate monolingual BIMOLA (i.e. French-only for the English word, English-only for the French word), which is shown in the two panels at the bottom of the figure. The situation amounts to a French speaker who does not understand a word of English, or an English speaker who does not comprehend one word of French. The model does activates candidates in its single language (e.g. F “soit”, “serve”, “signifie” for E “sing”; E “take”, “die”, “dash” for F “tâche”) but cannot find a word among them that matches the input, and so the model ends up dropping them all (exactly as when the model encounters a nonword). Taking a final global look at the figure, we note that the four panels on the left basically mirror the four panels on the right, with the roles of the two languages exchanged (although, of course, the details in the panels differ as they depend on the specific words).

From here on, we will always use BIMOLA as bilingual model and will at all times keep French as the base language (at 100%) and English as the guest language (at 80% language activation level), which is the default when BIMOLA

---

[44] The guest language candidates (e.g. E “sin”, “seem”, “thin” for E “sing”; F “passe” and “taise” for F “tâche” in the third row of panels) are not necessarily the same candidates as in monolingual BIMOLA (top row of panels). The reason for this is the attenuated lateral inhibition for guest language phonemes (see p. 88f).

is newly launched. We will examine more closely the recognition of English guest words in French (this corresponds, in Figure 23, to the left panel in the third line) and will study specific guest word properties and their effects.

Effects in bilingual guest word recognition. Bilinguals dynamically can bring in guest words into the base language in a variety of ways (see p. 87f). In code-switching, the guest word is pronounced using phonemes from the guest language (i.e. with English phonemes in the case of English guest words). But in borrowing, the guest word (still an English word) is phonetically adapted and pronounced with phonemes from the base language (i.e. French phonemes in the situation considered here). Of the factors that were shown, in experiments, to be involved in guest word recognition, three are a consequence of inherent properties of the guest words (Grosjean, 1988, 1997, 2008; Li, 1996; Schulpen et al., 2003; Lagrou et al., 2011). First, there is an effect of language phonetics: code-switches (guest words pronounced in the guest language) are recognized more easily and faster than borrowings (guest words pronounced in the base language). Second, if guest words are marked by their phonotactics (e.g. initial phonemes) to belong to the guest language rather than the base language, they are recognized more quickly than guest words not marked in such a way (effect of phonotactics). Third, guest words that have very similar sounding words, that is to say, near-homophones, in the base language, are recognized with more difficulty and more delay (or even not at all), compared to guest words that do not have near-homophones in the base language (effect of near-homophony).

The experiment by Grosjean (1988), which we employed and followed to evaluate BIMOLA, used the French carrier sentence “Il faudrait qu’on…” (“We should…”), completed either with a French verb (a filler) in the subjunctive[45] or with an English verb (a guest word), and ending with a French noun phrase that provided semantic context. For example, “Il faudrait qu’on slash tous les prix”, “Il faudrait qu’on lean contre le mur”, and “Il faudrait qu’on pick les bons chiffres”, were some of the experiment’s mixed-language utterances. Grosjean used 24 English guest words (all monosyllabic), 8 in each of the following three groups. Type 1 words, such as “slash”, were marked phonotactically to belong to the guest language: they were four-phoneme CCVC words whose initial CC (i.e. the /sl/ of “slash”) are frequent in English but infrequent in French. Type 2 words, like “lean”, were not marked phonotactically to belong to the guest language: they were three-phoneme CVC words whose initial CV (i.e. the /li/ of “lean”) are

[45] This is why BIMOLA has all its French verbs in the 3rd-person singular subjunctive.

common in French but uncommon in English. Type 3 words, for example, “pick”, were phonotactically much the same as Type 2 (three-phoneme CVC words with an initial CV that is more frequent in French than in English) but, in addition, they had a near-homophone in French, for instance, “pique” in the case of “pick”. Grosjean tested each guest word both as a code-switch (English pronunciation) and as a borrowing (French pronunciation). In the case of Type 3 words, the borrowing versions were absolutely indistinguishable from their counterparts in French (i.e. “pick” as borrowing was pronounced F /pik/, exactly as “pique”).

A set of 3 × 8 guest words is too small a sampling size to test a model: BIMOLA produces only one result per stimulus whereas the experiment had a response from several participants. Therefore, we extended Grosjean’s (1988) set of stimuli to 3 × 32 guest words (i.e. 32 Type 1 words, 32 Type 2 words, and 32 Type 3 words). Of course, only the guest word itself was tested in BIMOLA, not a whole sentence with context. But, as in Grosjean’s experiment, each of the guest words was run as a code-switch (i.e. with English phonemes, just as the word is stored in BIMOLA’s English lexicon) and also as a borrowing (with French phonemes). For the stimuli from Grosjean, we used the same borrowing pronunciation as in the experiment, and for the new stimuli, we established a suitably adapted pronunciation (e.g. F /slœm/ for E “slum”, F /kap/ for E “cap”, F /ʁɔk/ for E “rock”), and stored it in the stimulus list along with the English word. While selecting our additional stimuli, we carefully controlled the essential word properties within the bounds of BIMOLA’s lexicon. The three groups of stimuli do not differ in the frequency: means of 0.229 (*SD* = 0.154) for Type 1, 0.228 (*SD* = 0.189) for Type 2, and 0.229 (*SD* = 0.189) for Type 3. They do not differ in the uniqueness point: in each group, 11 words have a UP after their end, and 20 to 21 words have a UP on the last phoneme; one non-significant exception is “drop”, a Type 1 word (from Grosjean) with a UP on the vowel. Obviously, the length is not the same for Type 1 words (four phonemes) as for the Type 2 and Type 3 words (three phonemes), but by using the relative measure of %IP in percentage of word length including silence, length will be duly neutralized. For each stimulus, we determined the number of English words that start with the first two phonemes of the code-switch version as well as the number of French words that begin with the first two phonemes of the borrowing version, and so established, following Grosjean, an English to French ratio for that stimulus.[46]

[46] Two examples: for “swim”, 27 English words but only 3 French words start with /sw/ (hence its E to F ratio is 27 : 3, or 9 : 1), and for “cap”, 30 English words begin with /k^{h}æ/ and 88 French words start with /ka/ (so, its E to F ratio is 30 : 88, or 1 : 2.9).

For the Type 1 words, it ranges from 1.7 : 1 to 29 : 1, with a mean of 8.8 : 1. This signifies that all Type 1 words have a CC beginning that is favored by English. For the Type 2 words, the E to F ratio ranges from 1 : 1 to 1 : 7, with a mean of 1 : 2.5, and for the Type 3 words, it ranges from 1 : 1 to 1 : 8.5, with a mean of 1 : 3.1. This expresses that all Type 2 and Type 3 words have a CV beginning that is favored by French or is neutral. Finally, Type 2 and Type 3 words diverge by the characteristic that all Type 3 words (but none of the Type 2 words) have near-homophone counterparts in French, just as in Grosjean's smaller set of stimuli.

We ran the 3 × 32 guest words in BIMOLA, once as code-switches and once as borrowings (192 items in all), and we obtained the following results. Globally, there was a recognition success rate of 72.4%: 139 of 192 items were isolated by BIMOLA (i.e. the targets surpassed all other words, both English and French, before the end of the silence input and they remained at the top during the 16 extra cycles). This is a majority of the items. Breaking down into the word types, we found that 63 of 64 Type 1 items (98.4%), 58 of 64 Type 2 items (90.6%), but only 18 of 64 Type 3 items (28.1%) were successful. That is, almost all Type 1 and Type 2 items (the former even more so than the latter) but less than a third of the Type 3 items were isolated by BIMOLA. Splitting instead by the language phonetics, we observed that 79 of 96 code-switches (82.3%) and 60 of 96 borrowings (62.5%) were successful. That is, both code-switches and borrowings were isolated by BIMOLA, but code-switches better so than borrowings. In Table 12, we placed the BIMOLA results (*N* = 192 items) side by side with the results of Grosjean's (1988) experiment (*N* = 48 items), and divided them down by word type as well as by language phonetics. The

Table 12.
Guest word recognition success rate by type of word and language phonetics: experimental and BIMOLA results

Type of word	*Language phonetics*	*Grosjean (1988)*	*BIMOLA*
Type 1	Code-switch	93.8%	100.0%
	Borrowing	100.0%	96.9%
Type 2	Code-switch	89.6%	93.8%
	Borrowing	89.6%	87.5%
Type 3	Code-switch	52.0%	53.1%
	Borrowing	33.3%	3.1%

results bear a striking similarity (as one notices by comparing them line by line), with the exception of Type 3 borrowings (the last line). Yet, on closer inspection, one can see that in the experiment, language phonetics played less of a role for Type 1 and Type 2 words (Type 2 code-switches and borrowings were isolated equally well, and Type 1 code-switches were, oddly enough, isolated less well than borrowings) than it played for Type 3 words (where the difference between code-switches and borrowings was clear; cf. Grosjean, 1988, p. 246). Now, in BIMOLA, language phonetics seems to play a big role for all three word types (100% code-switches vs. 96.9% borrowings were isolated in the Type 1 words, 93.8% code-switches vs. 87.5% borrowings were isolated in the Type 2 words), and most importantly so for the Type 3 words (53.1% code-switches but only 3.1% borrowings isolated). A difference between code-switches and borrowings is very much appropriate in BIMOLA, especially so for the Type 3 words, where the borrowing version is undistinguishable from its near-homophone counterpart in French, which, as we will see, is usually most activated in its stead.[47]

In the further analysis, we followed Grosjean's example and looked into Type 1 and Type 2 words separately from Type 3 words. Indeed, as Grosjean remarked, Type 1 and Type 2 words conduct themselves rather differently from Type 3 words. The former are usually isolated by BIMOLA with ease whereas the latter are hard to isolate for BIMOLA or they are not isolated at all. For the Type 1 and Type 2 words alone, the relative measure of %IP (which, as always, provides us with the word isolation point, neutralized by length) was obtained, so as to be examined in an analysis of variance, with language phonetics as a within factor (code-switch vs. borrowing) and type of word as a between factor (Type 1 vs. Type 2). For 6 items (1 Type 1 borrowing, 2 Type 2 code-switches, and 3 Type 2 borrowings) that were isolated by BIMOLA during the additional cycles after the end of silence, we accepted a %IP larger than 100%. One item (a Type 2 borrowing) was not isolated at all[48]: we replaced its missing %IP with

---

[47] BIMOLA manifests all the differences between variables in the directions expected, even more clearly than the experiment did. Regarding Type 3 borrowings, we point out that Grosjean's participants heard ¾ of stimuli in English (i.e. test words) vs. ¼ of stimuli in French (filler words) and so were probably led to propose the English targets often. For a Type 3 borrowing in BIMOLA, the English target is very active too but has hardly ever a chance to beat the more active French near-homophone.

[48] For "sip" run as a borrowing (F /sip/), BIMOLA proposes "seep" (or "seek") instead of "sip". This is because French /i/ is closer to the /i/ of "seep" than to the /ɪ/ of "sip" in terms of distances in the bilingual phoneme space.

the grand mean (unweighted average of cell means; Keppel, 1991) of 78.155%. In the ANOVA, a main effect was found for the factor of language phonetics, *F*(1, 62) = 49.811, *MSE* = 54.025, *p* < .001; the mean %IP was 73.501% (*SD* = 12.732%) for code-switches and 82.671% (*SD* = 12.732%) for borrowings. So, code-switches are isolated by BIMOLA more quickly than borrowings. A main effect was also obtained for the factor of word type, *F*(1, 62) = 6.897, *MSE* = 247.614, *p* = .011; the %IP was on the average 74.434% (*SD* = 12.076%) for Type 1 and 81.739% (*SD* = 13.928%) for Type 2. That is, Type 1 words, which are marked phonotactically to belong to English, are isolated by BIMOLA earlier than Type 2 words, which are not marked in that way (effect of phonotactics). There was no interaction between the two factors (*F* < 1); the cell means (and *SD*'s) were: Type 1 code-switches: 70.195% (11.022%), Type 2 code-switches: 76.807% (13.616%), Type 1 borrowings: 78.672% (11.739%), and Type 2 borrowings: 86.671% (12.593%). The two effects simulated are independent of each other.

Figure 24 demonstrates two examples of the guest words we tested: a Type 1 word, "blush" (shown in the upper row), and a Type 2 word, "pat" (in the lower row), either word once run as a code-switch (panel on the left) and once run as a borrowing (panel on the right). For the code-switches, the guest words were pronounced using English phonemes, /blʌʃ/ and /pʰæt/ (written in blue in the figure), exactly as the two words are stored in BIMOLA's English lexicon. For the borrowings, the guest words were pronounced with French phonemes, /blœʃ/ and /pat/ (written in red in the figure), the adaptation we used of these English words to the French base language. As one can observe in the figure, the code-switching version of "blush" (E /blʌʃ/) is isolated faster (IP@cycle = 54) than the borrowing version of "blush" (F /blœʃ/, IP@cycle = 65), and likewise, the code-switched "pat" (E /pʰæt/) is isolated more rapidly (IP@cycle = 56) than the borrowed "pat" (F /pat/, IP@cycle = 68). This is an example of the language phonetics effect. Before comparing Type 1 word "blush" with Type 2 word "pat", we point out that the two words match in their frequency (value in the lexicon = 0.235) and uniqueness point (essentially on the last phoneme as "patronize" is of a lower frequency and has no influence on "pat"). However, "blush" and "pat" differ in their phonotactics: the beginning /bl/ is more frequent in English than in French (43 English words vs. 18 French words, an E to F ratio of 2.4 : 1, which is characteristic of a Type 1 word), whereas the beginning /pa/ is more common in French compared to the beginning /pʰæ/ in English (23 English words vs. 62 French words, E to F ratio of 1 : 2.7, typical of a Type 2 word). (No French verb is pronounced /blœʃ/ nor /pat/—the nouns "patte" and "pâte", both /pat/, are not

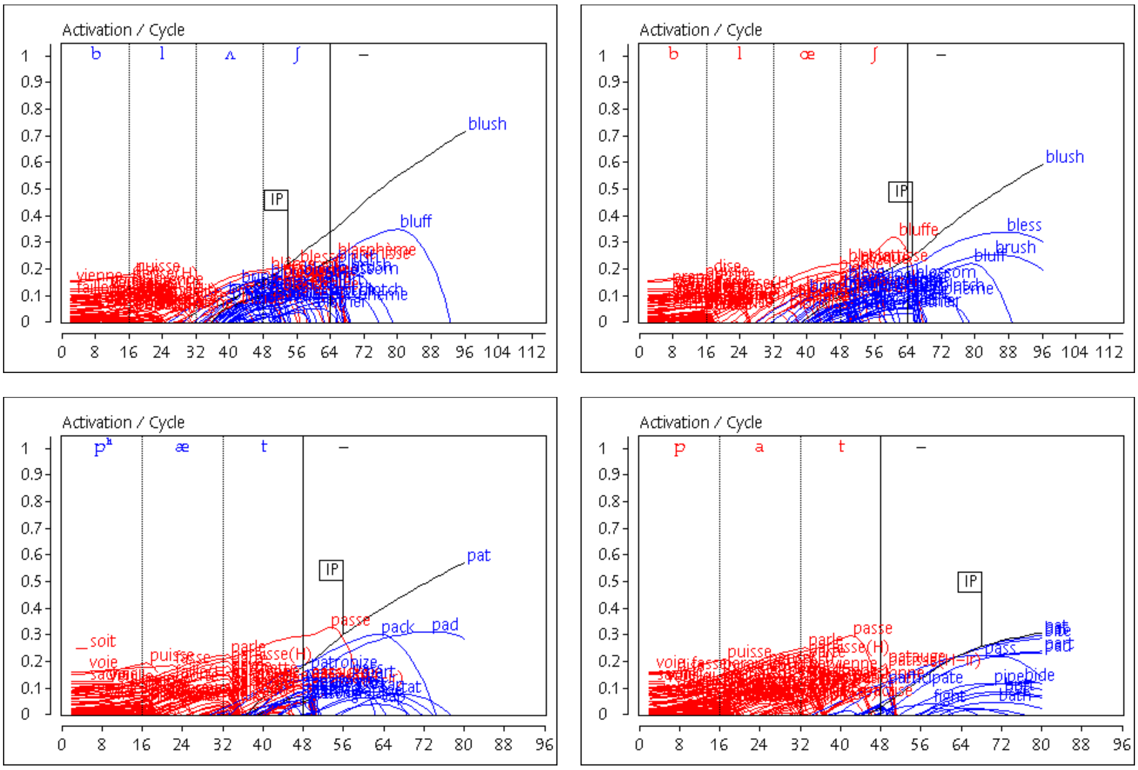


Figure 24. Running “blush”, a Type 1 word, and “pat”, a Type 2 word, one time as code-switches (on the left) and the other time as borrowings (on the right).

part of BIMOLA’s verb lexicon—and thus neither guest word is of the Type 3.) Comparing Type 1 word “blush” with Type 2 word “pat”, we find that the %IP is lower for “blush” than it is for “pat”: 67.50% vs. 87.50% in the case of the code-switches, and 81.25% vs. 106.25% in the case of the borrowings. It means that the portion of the guest word that is needed for its effective isolation in BIMOLA is a smaller one for “blush” (Type 1) than for “pat” (Type 2). This is an example of the phonotactics effect.

Let us finally turn our attention to the Type 3 words, which possess the phonotactic properties of the Type 2 words (i.e. a CV beginning that is favored by French or is neutral), but which, moreover, have French near-homophones. As we have seen in the general bilingual recognition study, guest words that have cross-language near-homophones (i.e. that are pronounced very similarly across English and French, such as E “clew” and F “cloue”, or E “boom” and F “boom”) are sometimes isolated by BIMOLA and sometimes they are not. Since BIMOLA lacks any help from semantic context, the isolation depends on three factors: whether the English or the French word is the more frequent one of the pair, whether English or French is the base language (here it is French), and whether the guest word is run as a code-switch (pronounced in English) or as a borrowing (i.e. pronounced identically to, and hence indistinguishable from, its counterpart in French). For 20 of our 32 Type 3 words, the English word has a higher frequency value in BIMOLA's lexicon than its French near-homophone (e.g. E "get" is more frequent than F "guette"): 15 of 20 were isolated if they were tested as code-switches, but only 1 of 20 (E "mean") was isolated if they were tested as borrowings. For 10 of the 32 Type 3 words, the English word is less frequent than its French counterpart (e.g. E "cool" compared to F "coule"): 2 of 10 (including "cool") were isolated when they were tested as code-switches, and 0 of 10 was isolated when they were tested as borrowings. For 2 of the 32 Type 3 words, the cross-language near-homophones have the same frequency (namely 0, for E "pan", F "pane", E "van", and F "vanne"): neither was isolated whether they were tested as code-switches or as borrowings. In short, Type 3 words have the best chance to be isolated by BIMOLA if they are run as code-switches and have a French near-homophone of lesser frequency; in the case of Type 3 words that are run as borrowings or have a French near-homophone of greater frequency, the near-homophones are most activated in their stead. As proposed and done by Grosjean (1988), we can subtract the Type 3 words' English frequency minus their French counterpart's frequency and so obtain an English frequency pull: the higher it is, the more the bilingual listener is pulled toward the English word; the lower it is (even negative when the French word is more frequent), the more the bilingual listener is drawn toward the French near-homophone. Is this also true in BIMOLA? For the 18 (15+1+2 of altogether 64) items successfully isolated by BIMOLA, a Pearson correlation coefficient r of −0.585 (p = .011) was found between the English frequency pull and the %IP. So, yes, the higher the English frequency pull, the earlier the model recognizes the Type 3 words.

Figure 25 presents two contrasting examples of Type 3 words: English "boot" (frequency of 0.131), more frequent than its French near-homophone "boute" (frequency of 0), in the upper row, and English "tat" (frequency of 0), less frequent than its French near-homophone "tâte" (frequency of 0.295), in the lower row. So, the pair "boot/boute" pulls toward English, and the pair "tat/tâte" pulls toward French. Both "boot" and "tat" are run one time as code-switches (left two panels), and the other time as borrowings and therefore phonetically indistinguishable from their French counterparts (right two panels). Of course, with French being the base language in the simulations here, BIMOLA activates predominantly French candidates, as one can see in all four panels, especially so at the start of the simulations. When "boot" is run as a code-switch (i.e. with

English phonemes /but/, see the top left panel), it becomes the best candidate during the /t/ (because the phonemes fed are English and because “boot” is more frequent than its French counterpart “boute”), and thus “boot” is isolated (IP@cycle = 47, %IP = 73.44%); French “boute” is activated too, as the second-best candidate. In the three other simulations (top right, bottom left and right), the French member of the near-homophone pair wins the race and the English member comes off second best. This happens either for the reason that the phonemes that are fed are French (top right: “boot” as a borrowing, pronounced F /but/, just like French “boute”), or because the English guest word is less frequent than its French counterpart (bottom left: “tat” as a code-switch, pronounced E /tʰæt/, in comparison with French “tâte”), or for both reasons together (bottom right: “tat” as a borrowing, pronounced F /tat/, now at an even larger disadvantage to “tâte”).

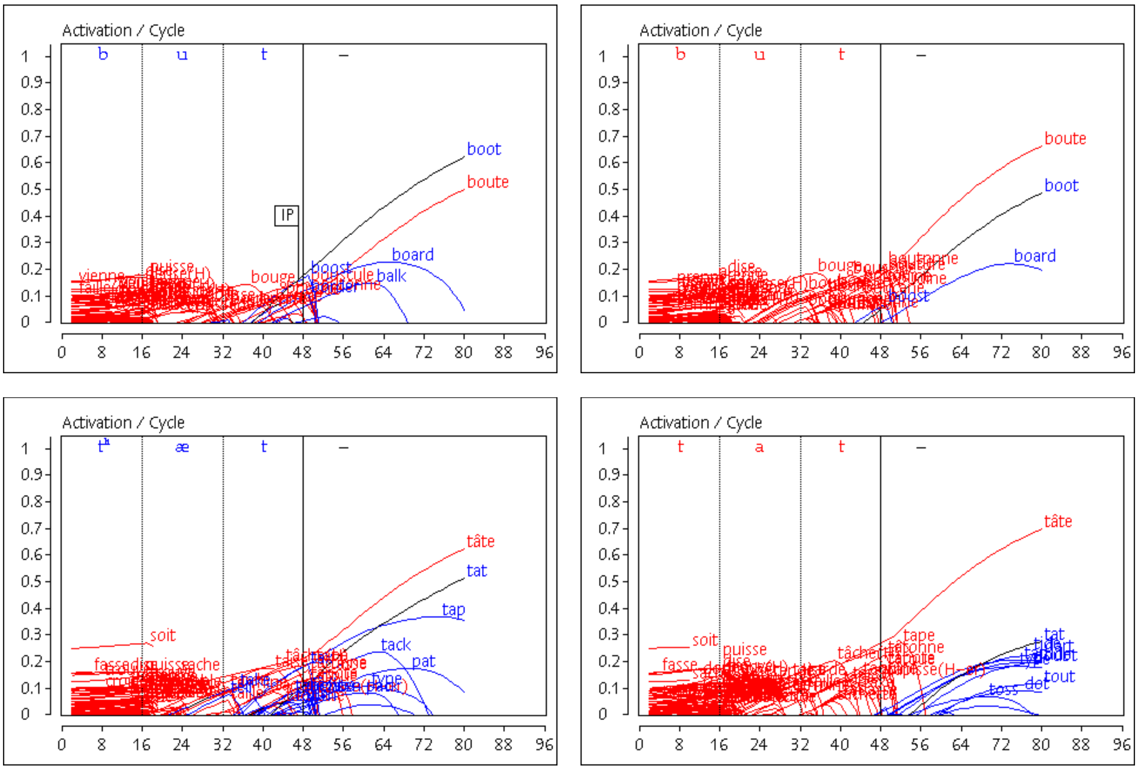


Figure 25. Running the Type 3 words “boot”, with frequency pull to English, and “tat”, with frequency pull to French, once as code-switches (left side) and once as borrowings (right side).

CHAPTER 8. CONCLUSION

The main contributions of this thesis are to describe and to thoroughly evaluate two new computer models of human spoken word recognition: the monolingual multiple-word recognition model FN5, and the bilingual single-word recognition model BIMOLA. Both models contain rich, detailed knowledge at three levels of linguistic description (words, phonemes, features). FN5's substantial lexicon of 17,668 French nouns, determiners, and prenominal adjectives, is available in a standard French and a Swiss French version (to put some dialectal differences into effect); it uses word frequency values based on a combination (and mutual correction) of two independent sources; and it includes pronunciation variants (schwa words; adjectives and determiners that take several word forms due to gender inflection, consonant liaison, or both; and words that can be contracted). BIMOLA's English–French lexicon of 8,696 verbs was prepared in such a way as to contain just the same number of words (4,348) in each language, and to have word frequency values that are comparable across the two languages. A phonetic feature matrix, covering all the phonemes of English and French (and including a few extra ones for Swiss French), is shared by the two models; it serves to quantify distances of phonemes (both within a language and between the languages) and to define a metric space of phonemes, which we visualized by means of hierarchical clustering analysis. Several internal mechanisms are common to both models: the activation and inhibition of phonemes (it can take place at various phoneme input delays and thus faster or slower speech rates); the activation and inhibition of words (the inhibition is caused by absent correct rather than present incorrect phonemes, the activation is influenced by word frequency and, if applicable, by variant frequency); and the isolation of words (with the isolation point as a useful measurement of recognition time, either in simulation cycles or as a percentage through the word). For either model, our evaluation approach was much the same: the evaluations comprised one part on the model's general performance (overall recognition success rate, and inspection of unsuccessful items) and another part on specific experimental effects (from the psycholinguistic literature) that are simulated by that model. Macros were set up, and used to batch process and efficiently run these many tests. Their results were analyzed using traditional statistical techniques and were amply illustrated with simulation examples.

The monolingual model FN5 is one of the first, if not the first, to simulate spoken word recognition in French (most previous psycholinguistic models were concerned with English). The internal mechanisms that are unique to FN5 are all related to sequential processing: they optimize and determine each word's position within a sequence of words (with the help of a position processor); they select a pronunciation variant, should the word have several variants (by using groups of connections and by attenuating the lateral inhibition for endings); and they check the words preceding/following the word, and its syntactic context. As a result, FN5 has a great overall capability to recognize single words as well as sequences of multiple words (of course, provided the words to be recognized by the model are present in its lexicon). Words in sequences can be connected and do not require any pause or explicit boundary between words; rather, FN5 finds the word boundaries by itself as a by-product of activating and recognizing the words from sequences of phonemes. By letting the model take advantage of certain contextual constraints (the number of words and the lexical category of the words in case of a sequence of words), we were able to further increase its general recognition success rate, from 91.3% to 99.7% for single words and from 83.6% to 99.9% for two-word sequences (as tested on 1,000 arbitrarily selected items each), that is, we brought it close to 100%.

Six specific psycholinguistic effects were shown to be simulated by FN5. The first three are language independent, and the other three are particular to French: 1. high-frequency words are isolated faster than low-frequency words (effect of frequency); 2. short words are isolated more quickly than long words (effect of length); 3. words with an early uniqueness point are isolated more rapidly than words with a late one (effect of uniqueness point, present at the normal speech rate, and absent at the fast speech rate); 4. words ending in a short vowel are isolated faster than those terminating in a long vowel (effect of vowel duration, applicable in the Swiss French version, and inapplicable in the standard French version); 5. schwa words are isolated more rapidly with schwa than without (effect of schwa deletion, observed for schwa words with optional or prohibited deletion, not for those with mandatory deletion); and 6. words are isolated more slowly when they occur in a sequence linked with liaison than when they occur alone, but equally fast in a sequence linked without liaison (effect of linking with liaison).

The BIMOLA model of bilingual spoken word recognition accounts for the bilingual's ability to process words from either of two languages, English and French, to go in and out of various language modes (bilingual vs. monolingual), and to recognize guest words, with or without guest language pronunciation. BIMOLA has two language networks (one for English, one for French), which employ the same general mechanisms as FN5 (i.e. activation and inhibition of phonemes, and of words), and which work simultaneously and independently from each other (without any inhibition across languages). While one language (the base language) is always set to 100% global activation, the other language (the guest language) is usually at a lower level of global activation, say, at 80%; this was implemented by using higher and lower resting levels. Even though BIMOLA is primarily a bilingual not a monolingual model, it can function with only one language, English or French. Its general recognition success rate, on the whole of either monolingual lexicon (single words only), is a perfect 100%. Bilingual BIMOLA deals with the two languages, English and French, together, and is able to identify words from both languages, regardless of which is the base language and which is the guest language, English or French. We found that base language words are recognized in bilingual BIMOLA as successfully (99.95%) and as quickly as the same words are recognized in monolingual BIMOLA. Guest words are recognized nearly as well (97.06% in case of English guest words, 98.28% in case of French guest words) but they usually take a little bit longer; we call this the guest word effect.

Six specific psycholinguistic effects more were shown to be simulated by BIMOLA. The first three are monolingual (they were already tested for French in FN5 and are described above, and were now revisited for English in BIMOLA): the effects of, 1. word frequency, 2. word length, and 3. uniqueness point. The other three effects accounted for by BIMOLA are bilingual, and concern the recognition of guest words: 4. guest words pronounced in the guest language, that is, code-switches, are isolated faster than guest words pronounced in the base language, that is, borrowings (effect of language phonetics); 5. guest words marked by their initial phonemes as belonging to the guest language are isolated more rapidly than guest words not marked as belonging to the guest language (effect of phonotactics); and lastly, 6. guest words that have near-homophones in the base language are isolated with more difficulty or not at all (due to the absence of semantic context), compared to guest words that do not have near-homophones in the base language (effect of near-homophony).

Comparing our two models, we realize that FN5, as a monolingual model for French, functions with only one language but is able to recognize both single words and multiple words connected in a sequence, while BIMOLA can operate with one language (monolingual BIMOLA) or two languages (bilingual BIMOLA) but it recognizes single words exclusively. This situation comes as no surprise since each of the two models was meant to pay attention to some aspects of spoken word recognition and had to disregard many other aspects, including those the other model focused on. (A systematic comparison between FN5 and monolingual BIMOLA was presented.) Even so, it would be a logical (and, of course, exciting) next step to attempt to combine the two models into one: a bilingual model for multiple-word recognition. It would involve, roughly speaking, all general mechanisms (common to our two current models), FN5's specific mechanisms (related to sequential, positional, and contextual processing), and BIMOLA's specific mechanisms (concerning global language modes, language activation, etc.). The features, which are already shared by FN5 and BIMOLA, would be the same, and the phonemes English and French. At the level of words, we could just merge BIMOLA's lexicon of English and French verbs with FN5's lexicon of French nouns, determiners, and adjectives (word frequency values are compatible between the two models), but we would surely need to cover some more verb forms (in the indicative in French) and probably also want to prepare a new lexicon of English nouns, determiners, and adjectives. Sequences of connected words to be processed by the future (much broader) model would be perhaps in the order, verb + determiner/adjective + noun (to be achieved by adapting FN5's position processor), and could be single-language utterances (e.g. "find that bike", "trouve ce vélo") as well as mixed-language utterances (e.g. "find that <u>vélo</u>", "trouve ce <u>bike</u>").

FN5 and BIMOLA come with a graphical user interface (which can be switched between English and French), run on Apple Macintosh computers (OS X), and are simple to install and use. In a number of boxes within the text, we provided information on how to load and manipulate components of the models (lexicons, parameters, position processor in the case of FN5, language mode for BIMOLA, etc.); examine the role of important new mechanisms; run a list of words (and, in FN5, of sequences of words) and get their isolation point; or test guest words with or without guest language pronunciation (in BIMOLA). As a result, parties interested in doing simulations with FN5 or BIMOLA on their own (e.g. in context of new experimental research), or using one or both models for teaching (e.g. with live demonstrations), should be able to do that with no trouble at all.

REFERENCES

Allopenna, P., Magnuson, J., & Tanenhaus, M. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439.

Altmann, G. (Ed.). (1990). *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

Altmann, G. (Ed.). (1997). *Cognitive models of speech processing: Psycholinguistic and computational perspectives on the lexicon*. Hove, UK: Psychology Press.

Auer, E., & Luce, P. (2005). Probabilistic phonotactics in spoken word recognition. In D. Pisoni & R. Remez (Eds.), *Handbook of speech perception* (pp. 610–630). Malden, MA: Blackwell.

Bagley, W. (1900). The apperception of the spoken sentence: A study in the psychology of language. *American Journal of Psychology*, *12*(1), 80–130.

Bard, E., & Shillcock, R. (1993). Competitor effects during lexical access: Chasing Zipf's tail. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 235–275). Hove, UK: Erlbaum.

Bard, E., Shillcock, R., & Altmann, G. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context.

Perception and Psychophysics, *44*(5), 395–408.
Berg, T., & Schade, U. (1992). The role of inhibition in a spreading-activation model of language production: I. The psycholinguistic perspective. *Journal of Psycholinguistic Research*, *21*(6), 405–434.
Berg, T., & Schade, U. (2000). A local connectionist account of consonant harmony in child language. *Cognitive Science*, *24*(1), 123–149.
Blumenfeld, H., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, *22*(5), 633–660.
Blumenfeld, H., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology*, *25*(5), 547–567.
Boden, M. (1988). *Computer models of mind: Computational approaches in theoretical psychology*. Cambridge, UK: Cambridge University Press.
Bronstein, I., & Semendjajew, K. (1991). *Taschenbuch der Mathematik: Ergänzende Kapitel* (6th ed.). Thun, Switzerland: Harri Deutsch.
Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.
Bürki, A., & Frauenfelder, U. (2012). Producing and recognizing words with two pronunciation variants: Evidence from novel schwa words. *Quarterly Journal of Experimental Psychology*, *65*(4), 796–824.
Calliope. (1989). *La parole et son traitement automatique*. Paris: Masson.
Cangelosi, A. (2005). The emergence of language: Neural and adaptive agent models. *Connection Science*, *17*(3/4), 185–190.
Chen, L. (2008). *Top-down effects on multiple meaning access within and between languages*. Doctoral dissertation, University of Michigan, Ann Arbor, MI.
Chomsky, N., & Halle, M. (1991). *The sound pattern of English*. Cambridge, MA: MIT Press. (Original work published 1968, New York: Harper & Row)
Christiansen, M., & Chater, N. (Eds.). (2001). *Connectionist psycholinguistics*. Westport, CT: Ablex.
Christophe, A., & Dupoux, E. (1996). Bootstrapping lexical acquisition: The role of prosodic structure. *Linguistic Review*, *13*(3/4), 383–412.
Cleland, A., Gaskell, G., Quinlan, P., & Tamminen, J. (2006). Frequency effects in spoken and visual word recognition: Evidence from dual-task

methodologies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 104–119.

Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.

Collins COBUILD English language dictionary. (1987). London: HarperCollins.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology: Section A*, *33*(4), 497–505.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Connine, C., & Pinnow, E. (2006). Phonological variation in spoken word recognition: Episodes and abstractions. *Linguistic Review*, *23*(3), 235–245.

Connine, C., Ranbom, L., & Patterson, D. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception and Psychophysics*, *70*(3), 403–411.

Connine, C., Titone, D., & Wang, J. (1993). Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 81–94.

Content, A., Mousty, P., & Radeau, M. (1990). Brulex: une base de données lexicales informatisée pour le français écrit et parlé. *L'Année psychologique*, *90*(4), 551–566.

Craig, C., & Kim, B. (1990). Effects of time gating and word length on isolated word-recognition performance. *Journal of Speech and Hearing Research*, *33*(4), 808–815.

Crocker, M. (1996). *Computational psycholinguistics: An interdisciplinary approach to the study of language*. Dordrecht, Netherlands: Kluwer Academic (Studies in Theoretical Psycholinguistics, Vol. 20).

Crocker, M., Pickering, M., & Clifton, C. (Eds.). (1999). *Architectures and mechanisms for language processing*. Cambridge, UK: Cambridge University Press.

Cutler, A. (1997). The comparative perspective on spoken-language processing. *Speech Communication*, *21*(1/2), 3–16.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 113–134). Hillsdale, NJ: Erlbaum.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 113–121.

Dahan, D., & Magnuson, J. (2006). Spoken word recognition. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 249–283). San Diego, CA: Academic Press.

Dahan, D., Magnuson, J., & Tanenhaus, M. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*(4), 317–367.

Davis, M., Marslen-Wilson, W., & Gaskell, G. (2002). Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 218–244.

de Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics*, *28*(4), 441–465.

de Bot, K., & Stoessel, S. (2000). In search of yesterday's words: Reactivating a long forgotten language. *Applied Linguistics*, *21*(3), 333–353.

Delattre, P. (1947). La liaison en français, tendances et classification. *French Review*, *21*(2), 148–157.

Delattre, P. (1951). Le jeu de l'e instable intérieur en français. *French Review*, *24*(4), 341–351.

Dell, F. (1985). *Les règles et les sons: introduction à la phonologie générative*. Paris: Hermann.

Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321.

Dell, G. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, *5*(4), 313–349.

Dijkstra, T., & de Smedt, K. (Eds.). (1996a). *Computational psycholinguistics: AI and connectionist models of human language processing*. London: Taylor & Francis.

Dijkstra, T., & de Smedt, K. (1996b). Computer models in psycholinguistics: An introduction. In T. Dijkstra & K. de Smedt (Eds.), *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 3–23). London: Taylor & Francis.

Dijkstra, T., & van Heuven, W. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Erlbaum.

Dijkstra, T., & van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197.

Dornbusch, T. (2012). *Orthographic influences on L2 auditory word processing*. Doctoral dissertation, Technische Universität Dortmund, Germany.

Dufour, S., Brunellière, A., & Frauenfelder, U. (2013). Tracking the time course of word-frequency effects in auditory word recognition with event-related potentials. *Cognitive Science*, *37*(3), 489–507.

Ellis, R., & Humphreys, G. (1999). *Connectionist psychology: A text with readings*. Hove, UK: Psychology Press.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211.

Elman, J., & McClelland, J. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27*(2), 143–165.

Encrevé, P. (1983). La liaison sans enchaînement. *Actes de la recherche en sciences sociales*, *46*, 39–66.

Encrevé, P. (1988). *La liaison avec et sans enchaînement: phonologie tridimensionnelle et usages du français*. Paris: Seuil.

Engel, A., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5*(1), 16–25.

Feldman, J., & Ballard, D. (1982). Connectionist models and their properties. *Cognitive Science*, *6*(3), 205–254.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, *18*(1), 7–44.

Forster, K. (1976). Accessing the mental lexicon. In R. Wales & E. Walker (Eds.), *New approaches to language mechanisms: A collection of psycholinguistic studies* (pp. 257–287). Amsterdam: North-Holland.

Fougeron, C., & Smith, C. (1999). French. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 78–81). Cambridge, UK: Cambridge University Press.

Francis, W., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Frauenfelder, U. (1996). Computational models of spoken word recognition. In T. Dijkstra & K. de Smedt (Eds.), *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 114–138). London: Taylor & Francis.

Frauenfelder, U., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 50–86).

Cambridge, MA: MIT Press.

Frauenfelder, U., & Peeters, G. (1998). Simulating the time course of word recognition: An analysis of lexical competition in TRACE. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 101–146). Mahwah, NJ: Erlbaum.

Frauenfelder, U., Scholten, M., & Content, A. (2001). Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes*, *16*(5/6), 583–607.

Gaskell, G. (2007). Statistical and connectionist models of speech perception and word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 55–69). Oxford, UK: Oxford University Press.

Gaskell, G., & Marslen-Wilson, W. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*(5/6), 613–656.

Gaskell, G., & Marslen-Wilson, W. (1999). Ambiguity, competition, and blending in spoken word recognition. *Cognitive Science*, *23*(4), 439–462.

Gaskell, G., Spinelli, E., & Meunier, F. (2002). Perception of resyllabification in French. *Memory and Cognition*, *30*(5), 798–810.

Gilbert, S., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology*, *44*(3), 297–337.

Gimson, A. (1989). *An introduction to the pronunciation of English* (4th ed., revised by S. Ramsaran). London: Edward Arnold.

Goldstone, R. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 3–28.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: North-Holland (Advances in Psychology, Vol. 83).

Grainger, J., & Jacobs, A. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*(3), 518–565.

Grainger, J., & Jacobs, A. (Eds.). (1998a). *Localist connectionist approaches to human cognition*. Mahwah, NJ: Erlbaum.

Grainger, J., & Jacobs, A. (1998b). On localist connectionism and psychological science. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 1–38). Mahwah, NJ: Erlbaum.

Grainger, J., & van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 1–23). Hauppauge, NY: Nova Science.

Grevisse, M. (1990). *Précis de grammaire française* (29th ed.). Paris: Duculot.

Grevisse, M. (1993). *Le bon usage* (13th ed., revised by A. Goosse). Paris: Duculot.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, *28*(4), 267–283.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics*, *38*(4), 299–310.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes*, *3*(3), 233–274.

Grosjean, F. (1996). Gating. *Language and Cognitive Processes*, *11*(6), 597–604.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–254). Mahwah, NJ: Erlbaum.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131–149.

Grosjean, F. (2008). *Studying bilinguals*. Oxford, UK: Oxford University Press.

Grosjean, F., Carrard, S., Godio, C., Grosjean, L., & Dommergues, J.-Y. (2007). Long and short vowels in Swiss French: Their production and perception. *Journal of French Language Studies*, *17*(1), 1–19.

Grosjean, F., & Frauenfelder, U. (Eds.). (1997). *A guide to spoken word recognition paradigms*. Hove, UK: Psychology Press.

Grosjean, F., & Gee, J. (1987). Prosodic structure and spoken word recognition. *Cognition*, *25*(1/2), 135–155.

Grosjean, F., & Itzler, J. (1984). Can semantic constraint reduce the role of word frequency during spoken-word recognition? *Bulletin of the Psychonomic Society*, *22*(3), 180–182.

Grosjean, F., & Soares, C. (1986). Processing mixed language: Some preliminary findings. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives* (pp. 145–179). Hillsdale, NJ: Erlbaum.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, illusions. *Biological Cybernetics*, *23*(4), 187–202.

Grossberg, S., Boardman, I., & Cohen, M. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 481–503.

Grossberg, S., & Myers, C. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, *107*(4), 735–767.

Hansen, A. (1994). Etude du E caduc: stabilisation en cours et variations lexicales. *Journal of French Language Studies*, *4*(1), 25–54.

Hardcastle, W., & Hewlett, N. (Eds.). (1999). *Coarticulation: Theory, data, and techniques*. Cambridge, UK: Cambridge University Press.

Harley, T. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes*, *8*(3), 291–309.

Harm, M., & Seidenberg, M. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*(1), 1–31.

Henderson, L., Weighall, A., Brown, H., & Gaskell, G. (2013). Online lexical competition during spoken word recognition and word learning in children and adults. *Child Development*, *84*(5), 1668–1685.

Hernandez, A., Bates, E., & Avila, L. (1996). Processing across the language boundary: A cross-modal priming study of Spanish–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(4), 846–864.

Howes, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America*, *29*(2), 296–305.

Howes, D., & Solomon, R. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, *41*(6), 401–410.

Huang, X., Ariki, Y., & Jack, M. (1990). *Hidden Markov models for speech recognition*. Edinburgh, UK: Edinburgh University Press.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, *5*(3), 299–314.

Imbs, P. (1971). *Etudes statistiques sur le vocabulaire français. Dictionnaire des fréquences. Vocabulaire littéraire des XIXe et XXe siècles* (Centre de recherche pour un Trésor de la langue française, CNRS, Nancy). Paris: Librairie Marcel Didier.

Jacobs, A., & Grainger, J. (1992). Testing a semistochastic variant of the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1174–1188.

Jacobs, A., Rey, A., Ziegler, J., & Grainger, J. (1998). MROM-p: An interactive activation, multiple readout model of orthographic and phonological processes in visual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 147–188). Mahwah, NJ: Erlbaum.

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.

Ju, M., & Luce, P. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, *15*(5), 314–317.

Jusczyk, P., & Luce, P. (2002). Speech perception and spoken word recognition: Past and present. *Ear and Hearing*, *23*(1), 2–40.

Kaufman, L., & Rousseeuw, P. (1990). *Finding groups in data: An introduction to cluster analysis*. New York: John Wiley & Sons.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Kirshenbaum, E. (2001). *Representing IPA phonetics in ASCII*. Unpublished manuscript, Hewlett-Packard Laboratories, Palo Alto, CA.

Klatt, D. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, *7*(3), 279–312.

Knecht, P. (1985). La Suisse romande. In R. Schläpfer (Ed.), *La Suisse aux quatre langues* (pp. 125–169). Geneva, Switzerland: Zoé.

Kohonen, T. (1988). The "neural" phonetic typewriter. *Computer*, *21*(3), 11–22.

Krause, E. (1986). *Taxicab geometry: An adventure in non-Euclidean geometry*. New York: Dover. (Original work published 1975, Menlo Park, CA: Addison-Wesley)

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Ladefoged, P. (1993). *A course in phonetics* (3rd ed.). Fort Worth, TX: Harcourt Brace Jovanovich.

Ladefoged, P. (1999). American English. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 41–44). Cambridge, UK: Cambridge University Press.

Laenzlinger, C. (2005). French adjective ordering: Perspectives on DP-internal movement types. *Lingua*, *115*(5), 645–689.

Lagrou, E., Hartsuiker, R., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 952–965.

Lagrou, E., Hartsuiker, R., & Duyck, W. (2013). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition*, *16*(3), 508–517.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*(3), 245–294.

Le nouveau Bescherelle: l'art de conjuguer. (1966). Paris: Hatier.

Le Petit Robert: dictionnaire de la langue française. (1992). Paris: Dictionnaires Le Robert.

Lee, K.-F. (1989). *Automatic speech recognition: The development of the SPHINX system*. Boston: Kluwer Academic.

Léon, P. (1966). Apparition, maintien et chute du "E" caduc. *La Linguistique*, *2*(2), 111–122.

Léon, P. (1996). *Phonétisme et prononciations du français*. Paris: Nathan.

Léwy, N., & Grosjean, F. (2008). The Léwy and Grosjean BIMOLA model. In F. Grosjean, *Studying bilinguals* (pp. 201–210). Oxford, UK: Oxford University Press.

Léwy, N., Grosjean, F., Grosjean, L., Racine, I., & Yersin, C. (2005). Un modèle psycholinguistique informatique de la reconnaissance des mots dans la chaîne parlée du français. *Journal of French Language Studies*, *15*(1), 25–48.

Li, P. (1996). Spoken word recognition of code-switched words by Chinese–English bilinguals. *Journal of Memory and Language*, *35*(6), 757–774.

Li, P., & Yip, M. (1998). Context effects and the processing of spoken homophones. *Reading and Writing*, *10*(3–5), 223–243.

Libben, G., & Jarema, G. (2002). Mental lexicon research in the new millennium. *Brain and Language*, *81*(1–3), 2–11.

Liberman, A. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.

Lippmann, R. (1989). Review of neural networks for speech recognition. *Neural Computation*, *1*(1), 1–38.

Lippmann, R. (1997). Speech recognition by machines and humans. *Speech Communication*, *22*(1), 1–15.

Luce, P., Goldinger, S., Auer, E., & Vitevitch, M. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics*, *62*(3), 615–625.

Luce, P., & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36.

Luce, P., Pisoni, D., & Goldinger, S. (1990). Similarity neighborhoods of spoken words. In G. Altmann (Ed.), *Cognitive models of speech processing:*

Psycholinguistic and computational perspectives (pp. 122–147). Cambridge, MA: MIT Press.

Magnuson, J., McMurray, B., Tanenhaus, M., & Aslin, R. (2003). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science*, *27*(2), 285–298.

Magnuson, J., Tanenhaus, M., & Aslin, R. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition*, *108*(3), 866–873.

Malécot, A. (1975). French liaison as a function of grammatical, phonetic and paralinguistic variables. *Phonetica*, *32*(3), 161–179.

Malécot, A. (1976). The effect of linguistic and paralinguistic variables on the elision of the French mute-e. *Phonetica*, *33*(2), 93–112.

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron*, *24*(1), 95–104.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115.

Marslen-Wilson, W. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 125–150). Hillsdale, NJ: Erlbaum.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1/2), 71–102.

Marslen-Wilson, W. (1993). Issues of process and representation in lexical access. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 187–210). Hove, UK: Erlbaum.

Marslen-Wilson, W., Moss, H., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1376–1392.

Marslen-Wilson, W., & Tyler, L. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*(4), 653–675.

Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63.

Massaro, D. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, *21*(3), 398–421.

Massaro, D. (1996). Modelling multiple influences in speech perception. In T. Dijkstra & K. de Smedt (Eds.), *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 85–113). London: Taylor & Francis.

Matthews, P. (1991). *Morphology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Mattys, S. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin and Review*, *4*(3), 310–329.

Mattys, S., Melhorn, J., & White, L. (2007). Effects of syntactic expectations on speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 960–977.

McClelland, J. (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitive Science*, *9*(1), 113–146.

McClelland, J. (1986). The programmable blackboard model of reading. In J. McClelland, D. Rumelhart, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models* (pp. 122–169). Cambridge, MA: MIT Press.

McClelland, J. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology*, *23*(1), 1–44.

McClelland, J. (2013). Integrating probabilistic models of perception and interactive neural networks: A historical and tutorial review. *Frontiers in Psychology*, *4*, 503.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McClelland, J., Mirman, D., Bolger, D., & Khaitan, P. (2014). Interactive activation and mutual constraint satisfaction in perception and cognition. *Cognitive Science*, *38*(6), 1139–1189.

McClelland, J., Mirman, D., & Holt, L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, *10*(8), 363–369.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*(5), 375–407.

McClelland, J., & Rumelhart, D. (1988). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. Cambridge, MA: MIT Press.

McQueen, J. (2007). Eight questions about spoken word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 37–53). Oxford, UK: Oxford University Press.

McQueen, J., & Cutler, A. (Eds.). (2001). *Spoken word access processes*. Hove, UK: Psychology Press.

McRae, K., Spivey-Knowlton, M., & Tanenhaus, M. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*(3), 283–312.

Merriam-Webster's collegiate dictionary (9th ed.). (1990). Springfield, MA: Merriam-Webster.

Métral, J.-P. (1977). Le vocalisme du français en Suisse romande: considérations phonologiques. *Cahiers Ferdinand de Saussure*, *31*, 145–176.

Metsala, J. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory and Cognition*, *25*(1), 47–56.

Miller, G. (1957). Some effects of intermittent silence. *American Journal of Psychology*, *70*(2), 311–314.

Miller, G., & Newman, E. (1958). Tests of a statistical explanation of the rank-frequency relation for words in written English. *American Journal of Psychology*, *71*(1), 209–218.

Miller, J. (1990). Speech perception. In D. Osherson & H. Lasnik (Eds.), *An invitation to cognitive science. Vol. 1: Language* (1st ed., pp. 69–93). Cambridge, MA: MIT Press.

Mirman, D., McClelland, J., Holt, L., & Magnuson, J. (2008). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science*, *32*(2), 398–417.

Moore, B., Tyler, L., & Marslen-Wilson, W. (Eds.). (2009). *The perception of speech: From sound to meaning*. Oxford, UK: Oxford University Press.

Morin, Y.-C., & Kaye, J. (1982). The syntactic bases for French liaison. *Journal of Linguistics*, *18*(1), 291–330.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178.

Morton, J. (1970). A functional model for memory. In D. Norman (Ed.), *Models of human memory* (pp. 203–254). New York: Academic Press.

Moss, H., & Marslen-Wilson, W. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1254–1276.

Myers-Scotton, C. (1992). Comparing codeswitching and borrowing. *Journal of Multilingual and Multicultural Development*, *13*(1/2), 19–39.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford, UK: Oxford University Press.

New, B. (2006). *Lexique 3: une nouvelle base de données lexicales*. Paper presented at the 13e Conférence Annuelle sur le Traitement Automatique du Langage Naturel (TALN), Leuven, Belgium.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, and Computers*, *36*(3), 516–524.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur Internet: Lexique™. *L'Année psychologique*, *101*(3), 447–462.

Nguyen, N. (2001). Rôle de la coarticulation dans la reconnaissance des mots. *L'Année psychologique*, *101*(1), 125–154.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234.

Norris, D. (2005). How do computational models help us develop better theories? In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 331–346). Mahwah, NJ: Erlbaum.

Norris, D., & McQueen, J. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395.

Norris, D., McQueen, J., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1209–1228.

Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299–325.

Norris, D., McQueen, J., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238.

Norris, D., McQueen, J., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, *34*(3), 191–243.

Norris, D., McQueen, J., Cutler, A., Butterfield, S., & Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes*, *16*(5/6), 637–660.

Oden, G., & Massaro, D. (1978). Integration of featural information in speech perception. *Psychological Review*, *85*(3), 172–191.

O'Reilly, R., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

O'Rourke, T., & Holcomb, P. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology*, *60*(2/3), 121–150.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences*, *23*(4), 443–467.

Paradis, M. (1989). Bilingual and polyglot aphasia. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 2, pp. 117–140). Amsterdam: Elsevier.

Park, S., & Miller, K. (1988). Random number generators: Good ones are hard to find. *Communications of the ACM*, *31*(10), 1192–1201.

Pastizzo, M., & Carbone, R. (2007). Spoken word frequency counts based on 1.6 million words in American English. *Behavior Research Methods*, *39*(4), 1025–1028.

Pfaff, C. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, *55*(2), 291–318.

Pieraccini, R. (2012). *The voice in the machine: Building computers that understand speech*. Cambridge, MA: MIT Press.

Pisoni, D., & Remez, R. (Eds.). (2005). *Handbook of speech perception*. Malden, MA: Blackwell.

Pitt, M., & Samuel, A. (1995). Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology*, *29*(2), 149–188.

Pitt, M., & Samuel, A. (2006). Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1120–1135.

Plaut, D., McClelland, J., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115.

Plunkett, K., & Elman, J. (1997). *Exercises in rethinking innateness: A handbook for connectionist simulations*. Cambridge, MA: MIT Press.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics*, *18*(7/8), 581–618.

Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, *26*(1), 47–104.

Quinlan, P. (1992). *The Oxford psycholinguistic database*. Oxford, UK: Oxford University Press.

R Development Core Team. (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available: http://www.R-project.org

Rabiner, L., & Juang, B.-H. (1993). *Fundamentals of speech recognition*. Englewood Cliffs, NJ: Prentice Hall.

Racine, I. (2008). *Les effets de l'effacement du schwa sur la production et la perception de la parole en français*. Doctoral dissertation, Université de Genève, Switzerland.

Racine, I., & Grosjean, F. (2000). Influence de l'effacement du schwa sur la reconnaissance des mots en parole continue. *L'Année psychologique*, *100*(3), 393–417.

Racine, I., & Grosjean, F. (2005). Le coût de l'effacement du schwa lors de la reconnaissance des mots en français. *Canadian Journal of Experimental Psychology*, *59*(4), 240–254.

Radeau, M., Morais, J., Mousty, P., & Bertelson, P. (2000). The effect of speaking rate on the role of the uniqueness point in spoken word recognition. *Journal of Memory and Language*, *42*(3), 406–422.

Radeau, M., Mousty, P., & Bertelson, P. (1989). The effect of the uniqueness point in spoken-word recognition. *Psychological Research*, *51*(3), 123–128.

Raphael, L., Borden, G., & Harris, K. (2011). *Speech science primer: Physiology, acoustics, and perception of speech* (6th ed.). Philadelphia: Lippincott Williams & Wilkins.

Reilly, R., & Sharkey, N. (Eds.). (1992). *Connectionist approaches to natural language processing*. Hove, UK: Erlbaum.

Reiser, M., & Wirth, N. (1992). *Programming in Oberon: Steps beyond Pascal and Modula*. Wokingham, UK: Addison-Wesley.

Roach, P. (2004). British English: Received Pronunciation. *Journal of the International Phonetic Association*, *34*(2), 239–245.

Roelofs, A. (2005). From Popper to Lakatos: A case for cumulative computational modeling. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 313–330). Mahwah, NJ: Erlbaum.

Roller, S. (1979). *Tableaux de conjugaison française*. Geneva, Switzerland: Département de l'instruction publique.

Rosenzweig, M., & Postman, L. (1957). Intelligibility as a function of frequency of usage. *Journal of Experimental Psychology*, *54*(6), 412–422.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart, J. McClelland, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Rumelhart, D., & McClelland, J. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60–94.

Salasoo, A., & Pisoni, D. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*, *24*(2), 210–231.

Samuel, A., & Pitt, M. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*(2), 416–434.

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change*, *2*(1), 71–101.

Schade, U., & Berg, T. (1992). The role of inhibition in a spreading-activation model of language production: II. The simulational perspective. *Journal of Psycholinguistic Research*, *21*(6), 435–462.

Schane, S. (1968). *French phonology and morphology*. Cambridge, MA: MIT Press.

Scharenborg, O., Norris, D., ten Bosch, L., & McQueen, J. (2005). How should a speech recognizer work? *Cognitive Science*, *29*(6), 867–918.

Schmid, M. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, *13*(1), 1–7.

Schulpen, B., Dijkstra, T., Schriefers, H., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(6), 1155–1178.

Seidenberg, M., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568.

Selfridge, O. (1959). Pandemonium: A paradigm for learning. In D. Blake & A. Uttley (Eds.), *Mechanisation of thought processes: Proceedings of a symposium held at the National Physical Laboratory* (Vol. 1, pp. 511–531). London: Her Majesty's Stationery Office.

Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences*, *16*(3), 417–451.

Shepard, R. (1980, October 24). Multidimensional scaling, tree-fitting, and clustering. *Science*, *210*(4468), 390–398.

Shillcock, R., Lindsey, G., Levy, J., & Chater, N. (1992). A phonologically motivated input representation for the modelling of auditory word perception in continuous speech. In J. Kruschke (Ed.), *Proceedings of the 14th Annual Conference of the Cognitive Science Society* (pp. 408–413). Hillsdale, NJ: Erlbaum.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory and Cognition*, *12*(4), 380–386.

Spinelli, E., McQueen, J., & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, *48*(2), 233–254.

Stemberger, J. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–186). London: Erlbaum.

Strauss, T., Harris, H., & Magnuson, J. (2007). jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods*, *39*(1), 19–30.

Struyf, A., Hubert, M., & Rousseeuw, P. (1996). Clustering in an object-oriented environment. *Journal of Statistical Software*, *1*(4), 1–30.

Struyf, A., Hubert, M., & Rousseeuw, P. (1997). Integrating robust clustering techniques in S-PLUS. *Computational Statistics and Data Analysis*, *26*(1), 17–37.

Taft, M., & Hambly, G. (1986). Exploring the Cohort model of spoken word recognition. *Cognition*, *22*(3), 259–282.

Thibault, A., & Knecht, P. (1997). *Dictionnaire suisse romand: particularités lexicales du français contemporain*. Geneva, Switzerland: Zoé.

Toscano, J., Anderson, N., & McMurray, B. (2013). Reconsidering the role of temporal order in spoken word recognition. *Psychonomic Bulletin and Review*, *20*(5), 981–987.

Tranel, B. (1987). *The sounds of French: An introduction*. Cambridge, UK: Cambridge University Press.

Tseng, J. (2003). Edge features and French liaison. In J.-B. Kim & S. Wechsler (Eds.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar* (pp. 313–333). Stanford, CA: CSLI.

Tyler, L. (1984). The structure of the initial cohort: Evidence from gating. *Perception and Psychophysics*, *36*(5), 417–427.

Tyler, L., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception and Psychophysics*, *34*(5), 409–420.

van Heuven, W., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483.

van Kuijk, D., Wittenburg, P., & Dijkstra, T. (1996). *RAW: A real-speech model for human word recognition*. Paper presented at the 4th International Conference on Spoken Language Processing, Philadelphia.

Waibel, A., & Lee, K.-F. (Eds.). (1990). *Readings in speech recognition*. San Mateo, CA: Morgan Kaufmann.

Walley, A., Michela, V., & Wood, D. (1995). The gating paradigm: Effects of presentation format on spoken word recognition by children and adults. *Perception and Psychophysics*, *57*(3), 343–351.

Walter, H. (1983). La nasale vélaire /ŋ/, un phonème du français? *Langue française*, *60*, 14–29.

Walter, H. (1999). French—an accommodating language: The chronology, typology and dynamics of borrowing. *Current Issues in Language and Society*, *6*(3/4), 195–220.

Waltz, D., & Pollack, J. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, *9*(1), 51–74.

Warren, P. (1993). *MacTRACE: An implementation on Macintosh of the TRACE model of spoken word recognition*. Unpublished manuscript, Department of Linguistics, University of Cambridge, Cambridge, UK.

Warren, P. (2013). Spoken word recognition. In *Introducing psycholinguistics* (pp. 119–138). Cambridge, UK: Cambridge University Press.

Wauquier-Gravelines, S. (1996). *Organisation phonologique et traitement de la parole continue: contraintes prosodiques et phonologiques de l'accès lexical*. Doctoral dissertation, Université Paris VII Diderot, Paris.

Wauquier-Gravelines, S. (1999). Segmentation lexicale de la parole continue: la linéarité en question. *Recherches linguistiques de Vincennes*, *28*, 133–156.

Wells, J. (1995). *Computer-coding the IPA: A proposed extension of SAMPA*. Unpublished manuscript, University College of London, London.

Wirth, N. (1986). *Algorithms and data structures*. Englewood Cliffs, NJ: Prentice Hall.

Yersin-Besson, C., & Grosjean, F. (1996). L'effet de l'enchaînement sur la reconnaissance des mots dans la parole continue. *L'Année psychologique*, *96*(1), 9–30.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*(1), 25–64.

APPENDICES

Appendix A.
The phonetic feature matrix

Language(s)*	Phoneme	Alloph.	Examples†	SONO	SYLL	CONS	COR	LOW	BACK	HIGH	FRO	ROU	NAS	LAT	CONT	DEL	TEN	VOI	ASP	LONG	MOV	Code
E	/p/	[p]	speak, opt, open, tap	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	.0	.0	1.	.0	.5	.1	.0	p
		[pʰ]	put, pile, appear, play	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	.0	.0	1.	.0	.7	.2	.0	P
F SF	/p/		poupée, lampe, pluie	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	.0	.0	.5	.0	.0	.0	.0	p
E	/b/		begin, bill, obey, barb	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	.0	.0	.5	1.	.5	.0	.0	b
F SF	/b/		bébé, balle, abeille	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	.0	.0	.0	1.	.0	.0	.0	b
E	/t/	[t]	steal, start, let, enter	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	.0	.0	1.	.0	.5	.1	.0	t
		[tʰ]	tell, try, retire, attend	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	.0	.0	1.	.0	.7	.2	.0	T
		[ɾ]	water, tidy, riddle	.0	.0	1.	1.	–	–	.5	.6	.0	.0	.0	.0	.0	.0	1.	.0	.1	.0	D
F SF	/t/		tarif, théâtre, petite	.0	.0	1.	1.	–	–	.5	.8	.0	.0	.0	.0	.0	.5	.0	.0	.0	.0	t
E	/d/		drive, deal, nod, undo	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	.0	.0	.5	1.	.5	.0	.0	d
F SF	/d/		droit, idée, splendide	.0	.0	1.	1.	–	–	.5	.8	.0	.0	.0	.0	.0	.0	1.	.0	.0	.0	d
E	/k/	[k]	skin, like, ask, react	.0	.0	1.	.0	–	–	1.	.2	.0	.0	.0	.0	.0	1.	.0	.5	.1	.0	k
		[kʰ]	clean, cry, keep, incur	.0	.0	1.	.0	–	–	1.	.2	.0	.0	.0	.0	.0	1.	.0	.7	.2	.0	K
F SF	/k/		colle, écho, sac, quel	.0	.0	1.	.0	–	–	1.	.2	.0	.0	.0	.0	.0	.5	.0	.0	.0	.0	k
E	/g/		get, agree, jog, giggle	.0	.0	1.	.0	–	–	1.	.2	.0	.0	.0	.0	.0	.5	1.	.5	.0	.0	g
F SF	/g/		goût, langue, regret	.0	.0	1.	.0	–	–	1.	.2	.0	.0	.0	.0	.0	.0	1.	.0	.0	.0	g
E	/f/		fit, flee, phone, suffer	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	1.	.0	1.	.0	.0	.8	.0	f
F SF	/f/		fils, effet, strophe	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	1.	.0	1.	.0	.0	.8	.0	f
E	/v/		voice, cover, give	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	1.	.0	1.	1.	.0	.7	.0	v
F SF	/v/		vrai, savon, veuve	.0	.0	1.	.0	–	–	.5	1.	.0	.0	.0	1.	.0	1.	1.	.0	.7	.0	v

E			/θ/		thank, froth, lengthen	.0	.0	1.	1.	–	–	.5	.8	.0	.0	.0	1.	.0	.5	.0	.0	.8	.0	Q
E			/ð/		smooth, bathe, gather	.0	.0	1.	1.	–	–	.5	.8	.0	.0	.0	1.	.0	.5	1.	.0	.7	.0	q
E			/s/		see, sell, concern, ice	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	1.	.0	1.	.0	.0	.8	.0	s
	F	SF	/s/		sol, garçon, cesse, os	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	1.	.0	1.	.0	.0	.8	.0	s
E			/z/		use, reason, zigzag	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	1.	.0	1.	1.	.0	.7	.0	z
	F	SF	/z/		zone, usine, gaz, aise	.0	.0	1.	1.	–	–	.5	.7	.0	.0	.0	1.	.0	1.	1.	.0	.7	.0	z
E			/ʃ/		show, finish, machine	.0	.0	1.	1.	–	–	1.	.5	.0	.0	.0	1.	.0	1.	.0	.0	.8	.0	S
	F	SF	/ʃ/		chat, chercheur, niche	.0	.0	1.	1.	–	–	1.	.5	.0	.0	.0	1.	.0	1.	.0	.0	.8	.0	S
E			/ʒ/		measure, sabotage	.0	.0	1.	1.	–	–	1.	.5	.0	.0	.0	1.	.0	1.	1.	.0	.7	.0	Z
	F	SF	/ʒ/		jeu, genou, âge, juge	.0	.0	1.	1.	–	–	1.	.5	.0	.0	.0	1.	.0	1.	1.	.0	.7	.0	Z
E			/ʧ/		cheer, watch, culture	.0	.0	1.	1.	–	–	1.	.7	.0	.0	.0	.0	1.	1.	.0	.0	.9	.0	C
E			/ʤ/		judge, adjust, stage	.0	.0	1.	1.	–	–	1.	.7	.0	.0	.0	.0	1.	1.	1.	.0	.8	.0	ç
E			/m/		mean, smile, murmur	1.	.0	1.	.0	–	–	.5	1.	.0	1.	.0	.0	.0	.5	1.	.0	.5	.0	m
	F	SF	/m/		miel, maman, âme	1.	.0	1.	.0	–	–	.5	1.	.0	1.	.0	.0	.0	.5	1.	.0	.5	.0	m
E			/n/		note, snow, run, end	1.	.0	1.	1.	–	–	.5	.7	.0	1.	.0	.0	.0	.5	1.	.0	.5	.0	n
	F	SF	/n/		nom, nid, nonne, une	1.	.0	1.	1.	–	–	.5	.8	.0	1.	.0	.0	.0	.5	1.	.0	.5	.0	n
	F	SF	/ɲ/		agneau, ligne, signal	1.	.0	1.	1.	–	–	.5	.4	.0	1.	.0	.0	.0	.5	1.	.0	.6	.0	G
	F	SF	/ŋ/		bowling, camping	1.	.0	1.	.0	–	–	.5	.2	.0	1.	.0	.0	.0	.5	1.	.0	.6	.0	N
E			/ŋ/		sing, finger, anchor	1.	.0	1.	.0	–	–	.5	.2	.0	1.	.0	.0	.0	.5	1.	.0	.6	.0	N
E			/l/	[l]	live, lay, blow, allow	1.	.0	1.	1.	–	–	.8	.7	.0	.0	1.	1.	.0	.5	1.	.0	.4	.0	l
				[ɫ]	call, file, help, pulse	1.	.0	1.	1.	–	–	.7	.6	.0	.0	1.	1.	.0	1.	1.	.0	.7	.0	L
	F	SF	/l/		lard, client, salle, le	1.	.0	1.	1.	–	–	.3	.8	.0	.0	1.	1.	.0	.5	1.	.0	.4	.0	l
E			/ɹ/		rain, progress, vary	1.	.0	1.	1.	–	–	.5	.6	.0	.0	.0	1.	.0	.5	1.	.0	.6	.0	r
	F	SF	/ʁ/		rare, prêt, arbre, or	1.	.0	1.	.0	–	–	.0	.0	.0	.0	.0	1.	.0	.5	1.	.0	.6	.0	R
E			/h/		hope, hunt, behave	1.	.0	.0	.0	–	–	.0	.0	.0	.0	.0	1.	.0	1.	.0	.0	.2	.0	h

E			/j/	yield, muse, eulogize	1.	.0	.0	.0	–	–	1.	.8	.0	.0	.0	1.	.0	.0	1.	.0	.5	.0	j
	F	SF	/j/	hier, bien, fille, paille	1.	.0	.0	.0	–	–	1.	.8	.0	.0	.0	1.	.0	.0	1.	.0	.5	.0	j
E			/w/	wet, wear, swim, quit	1.	.0	.0	.0	–	–	.7	.2	1.	.0	.0	1.	.0	.0	1.	.0	.6	.0	w
	F	SF	/w/	oui, boîte, noir, moi	1.	.0	.0	.0	–	–	.7	.2	1.	.0	.0	1.	.0	.0	1.	.0	.6	.0	w
	F	SF	/ɥ/	nuit, huile, duo, juin	1.	.0	.0	.0	–	–	.7	.7	1.	.0	.0	1.	.0	.0	1.	.0	.5	.0	W
	F	SF	/i/	lit, fini, difficile, ami	1.	1.	.0	.0	.0	.0	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.8	.2	i
		SF	/iː/	vie, amie, industrie	1.	1.	.0	.0	.0	.0	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.5	.2	i:
E			/i/	eat, meet, brief, key	1.	1.	.0	.0	.2	.1	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.5	i
E			/ɪ/	sit, wish, resist, itch	1.	1.	.0	.0	.3	.2	–	–	.0	.0	.0	1.	.0	.5	1.	.0	.7	.3	I
	F	SF	/y/	dû, jupe, luxe, rhume	1.	1.	.0	.0	.0	.3	–	–	1.	.0	.0	1.	.0	1.	1.	.0	.8	.2	y
		SF	/yː/	disparue, statue, vue	1.	1.	.0	.0	.0	.3	–	–	1.	.0	.0	1.	.0	1.	1.	.0	1.5	.2	y:
	F	SF	/e/	aîné, état, dernier, nez	1.	1.	.0	.0	.4	.1	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.9	.2	e
		SF	/eː/	allée, livrée, illuminée	1.	1.	.0	.0	.4	.1	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.5	.2	e:
E			/e/	set, head, lend, guess	1.	1.	.0	.0	.5	.2	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.7	.3	e
	F	SF	/ɛ/	lettre, fait, veine, vert	1.	1.	.0	.0	.7	.3	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.9	.2	E
		SF	/ɛː/	monnaie, paie, craie	1.	1.	.0	.0	.7	.3	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.5	.2	E:
	F	SF	/ø/	mieux, feu, chanteuse	1.	1.	.0	.0	.4	.4	–	–	1.	.0	.0	1.	.0	1.	1.	.0	.9	.2	ö
		SF	/øː/	queue, lieue, banlieue	1.	1.	.0	.0	.4	.4	–	–	1.	.0	.0	1.	.0	1.	1.	.0	1.5	.2	ö:
	F	SF	/œ/	fleur, sœur, seul, leur	1.	1.	.0	.0	.7	.5	–	–	.5	.0	.0	1.	.0	1.	1.	.0	.8	.2	Ö
E			/ɝ/	burn, circle, work, refer	1.	1.	.0	.0	.5	.6	–	–	.0	.0	.0	1.	.0	.5	1.	.0	1.	.5	ö
	F	SF	/ə/	degré, chemin, relevé	1.	1.	.0	.0	.5	.6	–	–	.0	.0	.0	1.	.0	.0	1.	.0	.6	.2	0
E			/ə/	apply, ripen, continue	1.	1.	.0	.0	.5	.6	–	–	.0	.0	.0	1.	.0	.0	1.	.0	.5	.3	0
E			/æ/	hand, add, ransack	1.	1.	.0	.0	1.	.3	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.7	.3	A
	F	SF	/a/	ananas, la, rat, choix	1.	1.	.0	.0	1.	.5	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.8	.2	a
		SF	/aː/	voie, joie, soie, proie	1.	1.	.0	.0	1.	.5	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.5	.2	a:

		SF	/ɑ/	grâce, tâche, dégât	1.	1.	.0	.0	1.	.6	–	–	.0	.0	.0	1.	.0	1.	1.	.0	.9	.2	A
E			/ɑ/	farm, star, calm, bath	1.	1.	.0	.0	1.	.7	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.5	a
E			/ʌ/	cut, jump, love, upset	1.	1.	.0	.0	.8	.6	–	–	.0	.0	.0	1.	.0	.5	1.	.0	.7	.3	V
E			/ɒ/	top, follow, offer, want	1.	1.	.0	.0	.9	.8	–	–	.2	.0	.0	1.	.0	1.	1.	.0	.7	.3	O
	F	SF	/ɔ/	folle, honorable, offre	1.	1.	.0	.0	.7	.7	–	–	.5	.0	.0	1.	.0	1.	1.	.0	.8	.2	o
E			/ɔ/	draw, force, cause	1.	1.	.0	.0	.5	.9	–	–	.5	.0	.0	1.	.0	1.	1.	.0	1.	.5	o
	F	SF	/o/	eau, gros, aube, rose	1.	1.	.0	.0	.4	1.	–	–	1.	.0	.0	1.	.0	1.	1.	.0	.9	.2	O
E			/ʊ/	push, look, wolf, bully	1.	1.	.0	.0	.3	.8	–	–	.2	.0	.0	1.	.0	.5	1.	.0	.7	.3	U
E			/u/	cool, move, rule, view	1.	1.	.0	.0	.2	.9	–	–	.5	.0	.0	1.	.0	1.	1.	.0	1.	.5	u
	F	SF	/u/	loup, sou, route, août	1.	1.	.0	.0	.0	1.	–	–	1.	.0	.0	1.	.0	1.	1.	.0	.8	.2	u
		SF	/uː/	roue, boue, hindoue	1.	1.	.0	.0	.0	1.	–	–	1.	.0	.0	1.	.0	1.	1.	.0	1.5	.2	u:
	F	SF	/ɛ̃/	fin, brin, main, singe	1.	1.	.0	.0	.8	.3	–	–	.0	1.	.0	1.	.0	1.	1.	.0	.9	.2	ê
		SF	/œ̃/	brun, aucun, humble	1.	1.	.0	.0	.8	.5	–	–	.5	1.	.0	1.	.0	1.	1.	.0	.9	.2	û
	F	SF	/ɑ̃/	danse, viande, sang	1.	1.	.0	.0	.8	.7	–	–	.0	1.	.0	1.	.0	1.	1.	.0	.9	.2	â
	F	SF	/ɔ̃/	ombre, compte, son	1.	1.	.0	.0	.8	1.	–	–	1.	1.	.0	1.	.0	1.	1.	.0	.9	.2	ô
E			/eɪ/	take, break, say, wait	1.	1.	.0	.0	.5	.2	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.6	1
E			/aɪ/	find, lie, light, fly, eye	1.	1.	.0	.0	1.	.5	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	1.	2
E			/ɔɪ/	join, point, enjoy, oil	1.	1.	.0	.0	.7	.9	–	–	.5	.0	.0	1.	.0	1.	1.	.0	1.	1.	3
E			/əʊ/	know, go, load, hold	1.	1.	.0	.0	.5	.6	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.6	4
E			/aʊ/	shout, vow, bounce	1.	1.	.0	.0	1.	.7	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.9	5
E			/ɪə/	hear, pierce, realize	1.	1.	.0	.0	.3	.2	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.8	6
E			/ɛə/	share, repair, bear	1.	1.	.0	.0	.7	.3	–	–	.0	.0	.0	1.	.0	1.	1.	.0	1.	.7	7
E			/ʊə/	assure, secure, tour	1.	1.	.0	.0	.3	.8	–	–	.2	.0	.0	1.	.0	1.	1.	.0	1.	.8	8

* English (E) and standard French (F) are used in BIMOLA; standard French and Swiss French (SF) are used in FN5.

† As far as the spelling to sound correspondence allows, the sound's approximate position is indicated by underlining.

Appendix B1.
Lexicon files that were prepared for BIMOLA

File	*Description*	*No. of words*
	Primary lexicon files	
BimolaVerbs.Lex	Bilingual verb lexicon, used in the bilingual simulations (combines the following two files)[a]	8,696
EnglishVerbs.Lex	English verb lexicon, used in the English monolingual simulations[a]	4,348
FrenchVerbs.Lex	French verb lexicon, used in the French monolingual simulations[a]	4,348
	Secondary lexicon files	
HomophEVerbs.Lex	English homophonous verbs of lower frequency	69
HomophFVerbs.Lex	French homophonous verbs of lower frequency	41
EnglishAuxiliaries.Lex	English auxiliaries	15
ArchaicEnglishVerbs.Lex	Some archaic English verbs	20
MoreEnglishVerbs.Lex	Additional English verbs (not integrated into EnglishVerbs.Lex)	1,105
AllEnglishVerbs.Lex	All English verbs (combining all English files)[a]	5,557
AllFrenchVerbs.Lex	All French verbs (combining all French files)[a]	4,389

[a] These lexicon files include the uniqueness point of each word (in relation to the other words in the same file).

Appendix B2.
Lexicon files that were prepared for FN5

All these files exist both in versions for Swiss French and standard French. To obtain the actual file names, substitute <Version> with “SwissFrench” or “French”, and replace <V> with “SF” or “F”, respectively.

File	*Description*	*No. of words*
	Primary lexicon files	
<Version>Ptitami.Lex	Nouns + adjectives and determiners, as pronounced before nouns (combining the following two files)	17,668
<Version>Nouns.Lex	Nouns	16,971
<Version>Ants.Lex	Adjectives and determiners, as pronounced before nouns[a]	697
<Version>IsoAnts.Lex	The same adjectives and determiners, as pronounced in isolation[a]	697
	Secondary lexicon files	
<V>.PlurAnts.Lex	Adjectives and determiners that are used before plural nouns[a]	34
<V>.IsoPlurAnts.Lex	The same words, as pronounced in isolation[a]	34
<V>.Alphabet.Lex	Nouns designating the letters of the alphabet	26
<V>.HomogrNouns.Lex	Nouns homographic with one of our adjectives or determiners	328

[a] For the files of adjectives and determiners, the versions “before nouns” and “in isolation” contain the same words and hence should not be loaded both at the same time.

Appendix C1.
Main macros used in the FN5 evaluations

Command	*Function*
Exec “InitFN5.French.Macro”	Load lexicon in standard French version and default parameters.
Exec “InitFN5.SwissFrench.Macro”	Load lexicon in Swiss French version and default parameters.
Exec “BatchFN5.Gen.Iso.Macro”	Examine general recognition of single isolated words.
Exec “BatchFN5.Iso.Freq.Macro”	Examine word frequency.
Exec “BatchFN5.Iso.Len.Macro”	Examine word length.
Exec “BatchFN5.Iso.UP.Macro”	Examine word uniqueness point at a normal speech rate.
Exec “BatchFN5.Radeau00.Macro”	Examine word uniqueness point at a fast speech rate.
Exec “BatchFN5.VDuration.Macro”	Examine vowel duration.
Exec “BatchFN5.Gen.Seq.Macro”	Examine general recognition of two connected words.
Exec “BatchFN5.Gen.SeqP.Macro”	Examine general recognition of two words separated by a pause.
Exec “BatchFN5.RG00.Macro”	Examine schwa deletion when it is optional.
Exec “BatchFN5.RG05.Macro”	Examine schwa deletion for mandatory vs. optional vs. prohibited.
Exec “BatchFN5.YBG96.Macro”	Examine linking with and without liaison (original list).
Exec “BatchFN5.YBG96a.Macro”	Examine linking with and without liaison (modified list).
Exec “BatchFN5.YBG96Ext.Macro”	Examine linking with and without liaison (extended list).

Appendix C2.
Main macros used in the BIMOLA evaluations

Command	*Function*
Exec “InitBimola.EOnly.Macro”	Load English lexicon, and set language mode to English-only.
Exec “InitBimola.FOnly.Macro”	Load French lexicon, and set language mode to French-only.
Exec “BatchBimola.Gen.EOnly.Macro”	Examine general monolingual recognition in English.
Exec “BatchBimola.Gen.FOnly.Macro”	Examine general monolingual recognition in French.
Exec “BatchBimola.FxLxU.Macro”	Examine three monolingual effects (word frequency, length, and uniqueness point).
Exec “InitBimola.Biling.Macro”	Load bilingual lexicon, and set language mode to BIMOLA’s default (French as base language and English as guest language at 80%).
Exec “BatchBimola.Gen.FBiling.Macro”	Examine general bilingual recognition with French as base language.
Exec “BatchBimola.Gen.EBiling.Macro”	Examine general bilingual recognition with English as base language.
Exec “BatchBimola.T123.Macro”	Examine guest word properties (effects of language phonetics, phonotactics, and near-homophony).
Exec “BatchBimola.LexOverlap.Macro”	Examine BIMOLA’s lexical overlap with FN5.
Exec “BatchFN5.LexOverlap.Macro”	Examine FN5’s lexical overlap with BIMOLA.

Appendix D1.
New stimuli used in the FN5 evaluations

Frequency

High accord, air, ami, amour, chambre, chose, coup, dieu, esprit, fille, fils, force, guerre, heure, histoire, homme, main, mal, moment, monde, mot, nuit, ordre, peu, peur, pouvoir, raison, temps, tête, travail.

Low algue, auge, braille, cancre, cèpe, crin, croupe, daim, drain, éclipse, effusion, établi, étole, faon, gond, gousse, guet, gui, jarre, latte, lobe, miche, mollet, motte, omble, paon, pus, quotient, rotin, truffe.

Length

Short aphte, chiot, chope, cuve, daube, deuil, dinde, douve, fugue, gauche, guêpe, gueule, gueuse, isthme, jade, jauge, jonque, juin, niais, nimbe, œuvre, ogre, ombre, oncle, quille, sud, tube, urne, veuve, vogue.

Long boutoir, braillard, brouillard, coloris, commandant, conscience, cornette, coursier, courtier, différend, divorce, faction, figurant, gestion, grillage, intendance, justice, martel, moustique, patience, patiente, portage, prieur, quartier, résident, révolte, savoir, sentiment, trafic, voyance.

Linking with and without liaison

With liaison léger~ ail (rail*), leurs~ ailes (zèle), mon~ air (nerf), petit~ ami (tamis), grand~ an (temps), long~ arçon (garçon), cent~ as (tasse), léger~ entier (rentier), premier~ épi (répit), huit~ ermites (termite), tout~ est (test), léger~ éveil (réveil), ancien~ hectare (nectar), brillant~ if (tif), cent~ ires (tir), son~ œuf (neuf), ton~ ombre (nombre), son~ once (nonce), cent~ urnes (turne), léger~ ut (rut).

Without liaison, C to V neuf ails (faille), chaque ancre (cancre), antique anneau (canot), fade anse (danse), même arbre (marbre), mille armes (larme), leur as (race), tragique aval (cavale), immonde épi (dépit), rapide essai (décès), ferme hélice (mélisse), dense heure (sœur), cette huile (tuile), sévère humeur (rumeur), flasque if (kif), même île (mil), macabre onde (ronde), neuf orges (forge), jeune os (noce), chaque ours (course).

Without liaison, C to C tragique lac (claque), neuf lames (flamme), vague lande (glande), neuf leurres (fleur), vague lobe (globe), neuf luths (flûte), sordide rame (drame), injuste rampe (trempe), solide rat (drap), modique rayon (crayon), vaste reine (traîne), vague rêne (graine), unique rétine (crétine), chaque rime (crime), tragique roi (croix), vif roman (froment), cette rousse (trousse), unique route (croûte), chaque rue (crue), triste ruelle (truelle).

* Potential distractors are given in parentheses.

Appendix D2.
New stimuli used in the BIMOLA evaluations

Frequency × Length × UP

Low, Short, Early	noise, ogle, okay, oust, thieve, usher, yacht, yearn.
Low, Long, Early	canvass, cement, dissuade, grimace, obtrude, reaffirm, scrounge, unbind.
High, Short, Early	ask, assure, choose, guide, honour, look, offer, view.
High, Long, Early	depend, direct, imagine, involve, maintain, mention, result, select.
Low, Short, Late	beam, coil, fib, hush, lout, nudge, rake, sneer.
Low, Long, Late	bandage, conserve, disobey, engrave, impute, inflate, overreach, unstitch.
High, Short, Late	file, guess, keep, laugh, lean, name, pull, teach.
High, Long, Late	conceive, concern, destroy, disappear, discuss, intend, reduce, refuse.

Guest words

Code-switch and borrowing pronunciations, and French near-homophones for Type 3, are given in parentheses. The first 8 of the 32 stimuli in each group are from Grosjean (1988).

Type 1 blot (/blɒt/ - F /blɔt/), drop (/dɹɒp/ - F /dʁɔp/), quit (/kʰwɪt/ - F /kwit/), skate (/skeɪt/ - F /sket/), skip (/skɪp/ - F /skip/), slash (/slæʃ/ - F /slaʃ/), snap (/snæp/ - F /snap/), sneak (/snik/ - F /snik/), bleed (/blid/ - F /blid/), blush (/blʌʃ/ - F /blœʃ/), clean (/kʰlin/ - F /klin/), dread (/dɹed/ - F /dʁɛd/), drill (/dɹɪɫ/ - F /dʁil/), muse (/mjuz/ - F /mjuz/), mute (/mjut/ - F /mjut/), quill (/kʰwɪɫ/ - F /kwil/), scan (/skæn/ - F /skan/), slack (/slæk/ - F /slak/), slum (/slʌm/ - F /slœm/), snuff (/snʌf/ - F /snœf/), speak (/spik/ - F /spik/), spin (/spɪn/ - F /spin/), spot (/spɒt/ - F /spɔt/), state (/steɪt/ - F /stet/), steal (/stiɫ/ - F /stil/), steep (/stip/ - F /stip/), step (/step/ - F /stɛp/), stun (/stʌn/ - F /stœn/), swell (/sweɫ/ - F /swɛl/), swim (/swɪm/ - F /swim/), swoon (/swun/ - F /swun/), twit (/tʰwɪt/ - F /twit/).

Type 2 beep (/bip/ - F /bip/), dab (/dæb/ - F /dab/), feed (/fid/ - F /fid/), lead(/i/) (/lid/ - F /lid/), lean (/lin/ - F /lin/), sip (/sɪp/ - F /sip/), tag (/tʰæg/ - F /tag/), tease (/tʰiz/ - F /tiz/), ban (/bæn/ - F /ban/), beam (/bim/ - F /bim/), beat (/bit/ - F /bit/), booze (/buz/ - F /buz/), cap (/kʰæp/ - F /kap/), con (/kʰɒn/ - F /kɔn/), give (/gɪv/ - F /giv/), gnash (/næʃ/ - F /naʃ/), lam (/læm/ - F /lam/), laugh (/lɑf/ - F /laf/), leave (/liv/ - F /liv/), man (/mæn/ - F /man/), map (/mæp/ - F /map/), move (/muv/ - F /muv/), nab (/næb/ - F /nab/), pad (/pʰæd/ - F /pad/), pat (/pʰæt/ - F /pat/), pool (/pʰuɫ/ - F /pul/), rag (/ɹæg/ - F /ʁag/), root (/ɹut/ - F /ʁut/), seal (/siɫ/ - F /sil/), tool (/tʰuɫ/ - F /tul/), toot (/tʰut/ - F /tut/), vat (/væt/ - F /vat/).

Type 3 cool (/kʰuɫ/ - F /kul/ coule), fool (/fuɫ/ - F /ful/ foule), knot (/nɒt/ - F /nɔt/ note), lease (/lis/ - F /lis/ lisse), peel (/pʰiɫ/ - F /pil/ pile), pick (/pʰɪk/ - F /pik/ pique), sit (/sɪt/ - F /sit/ cite), wrap (/ɹæp/ - F /ʁap/ râpe), bag (/bæg/ - F /bag/ bague), boot (/but/ - F /but/ boute), can (/kʰæn/ - F /kan/ cane), cash (/kʰæʃ/ - F /kaʃ/ cache), coop (/kʰup/ - F /kup/ coupe), fan (/fæn/ - F /fan/ fane), feel (/fiɫ/ - F /fil/ file), foot (/fʊt/ - F /fut/ foute), get (/get/ - F /gɛt/ guette), lash (/læʃ/ - F /laʃ/ lâche), mash (/mæʃ/ - F /maʃ/ mâche), mass (/mæs/ - F /mas/ masse), mat (/mæt/ - F /mat/ mate), mean (/min/ - F /min/ mine), mock (/mɒk/ - F /mɔk/ moque), nap (/næp/ - F /nap/ nappe), pan (/pʰæn/ - F /pan/ pane), peep (/pʰip/ - F /pip/ pipe), rock (/ɹɒk/ - F /ʁɔk/ roque), sap (/sæp/ - F /sap/ sape), sell (/seɫ/ - F /sɛl/ scelle), tat (/tʰæt/ - F /tat/ tâte), top (/tʰɒp/ - F /tɔp/ tope), van (/væn/ - F /van/ vanne).