

**Computational Psycholinguistics & Spoken Word
Recognition in The Bilingual & The Monolingual**
علم النفس اللغوي الحاسوبي والتعرف على الكلمات المنطوقة في ثنائي
اللغة وأحادي اللغة

Shaimma Abdulhameed Yousif

م.م شيماء عبدالحميد يوسف

Iraqi Ministry of Education

وزارة التربية العراقية

Maryam Kamil Kadhim

مريم كامل كاظم

University of Babylon /Faculty of Basic Education College

جامعة بابل كلية التربية الاساسية

bas858.mariem.kadhum@uobabylon.edu.iq



ABSTRACT

This is a study in computational psycholinguistics, an interdisciplinary research area that combines linguistics, psychology, and computer science knowledge and experiences. The simulation of human beings is the topic of the study. Word recognition, in other words, seeks to simulate on a machine the cognitive mechanism by which we enable and access words, their shapes, in our mental lexicon, as well as their constituent parts. It is concerned with the simulation of auditory (rather than written) word units since it is concerned with spoken (rather than written) word units.

This study, in particular, proposes two new models that we created and designed ourselves: The first, known as FN5, simulates spoken word comprehension in monolinguals, while the second, known as BIMOLA, models multilinguals. Bilinguals' spoken word comprehension. The monolingual model FN5 is based on French and thus includes a lexicon of 17,668 French words (nouns, determiners, and prenominal adjectives), some of which are prenominal adjectives. FN5 recognizes both single (isolated) words and sequences of two related words (determiner + noun or prenominal adjective + noun). It employs a novel approach to understanding sequences of words.

words by optimizing the alignment positions and pronunciation variants of the words. Furthermore, it allows for a variety of phonological phenomena to occur within a word or at its boundaries.

Key words: Bilinguals, monolingual, lexicon, computational.

هذه دراسة في اللغويات النفسية الحاسوبية، وهو مجال بحثي متعدد التخصصات يجمع بين المعرفة والخبرات اللغوية وعلم النفس وعلوم الكمبيوتر. ومحاكاة البشر هي موضوع الدراسة. بعبارة أخرى، يسعى التعرف على الكلمات إلى محاكاة الآلية المعرفية التي من خلالها نتمكن من الوصول إلى الكلمات وأشكالها في معجمنا العقلي، وكذلك أجزاءها المكونة، وهو يهتم بمحاكاة وحدات الكلمات السمعية (وليس المكتوبة) لأنه يهتم بوحدات الكلمات المنطوقة (وليس المكتوبة). تقترح هذه الدراسة، على وجه الخصوص، نموذجين جديدين قمنا بإنشائهما وتصميمهما بأنفسنا: الأول، المعروف باسم FN5، يحاكي فهم الكلمات المنطوقة لدى أحادي اللغة، بينما الثاني، المعروف باسم BIMOLA، يحاكي متعدد اللغات. فهم الكلمات المنطوقة لدى ثنائي اللغة يعتمد النموذج الأحادي اللغة FN5 على اللغة الفرنسية وبالتالي يتضمن معجمًا من ١٧٦٦٨ كلمة فرنسية (أسماء، ومحددات، وصفات ما قبل الاسم)، بعضها صفات ما قبل الاسم. يتعرف FN5 على الكلمات المفردة (المعزولة) وتسلسلات كلمتين مرتبطتين (محدد + اسم أو صفة ما قبل الاسم + اسم). وهو يستخدم نهجًا جديدًا لفهم تسلسلات الكلمات. الكلمات من خلال تحسين مواضع المحاذاة ومتغيرات النطق للكلمات وعلاوة على ذلك، فإنه يسمح لمجموعة متنوعة من الظواهر الصوتية بالحدوث داخل الكلمة أو عند حدودها.

الكلمات المفتاحية: ثنائي اللغة، أحادي اللغة، المعجم، الحاسوبي.



Aims and Objectives

One example is word recognition (or lexical access, as it is also known interchangeably). one of the most fundamental topics being researched in adult cognitive science psycholinguistics. For the auditory modality, it concerns the issue of how the The human mind is capable of perceiving and recalling individual spoken words. their real-time shape, be the words heard separately (i.e. separately), or be heard as part of a continuous flowing speech Experimentation The study of spoken word comprehension dates back over a century. already existed (see, for example, Bagley, 1900, for a very early study) and has since become a a vast and intense field with its own behavioral research, laboratory methods, and so on and paradigms (see Grosjea's guide).A variety of models of spoken word comprehension have been proposed in order to clarify and consolidate the many scientific findings that have accrued over time. Theorizing and posing were initially done verbally. That is, the pioneering early designers represented and illustrated the method in the form of a letter, even though this was often mixed.

Subject, Scope & Short Literature Review

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 2TRACE'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).

Originality of the Study

How is the thesis different from previous work on the topic? What is the contribution of this thesis to existing knowledge?) 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%.

References

1. 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.
2. Altmann, G. (Ed.). (1990). *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.
3. Altmann, G. (Ed.). (1997). *Cognitive models of speech processing: Psycholinguistic and computational perspectives on the lexicon*. Hove, UK: Psychology Press.
4. 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.
5. Bagley, W. (1900). The apperception of the spoken sentence: A study in the psychology of language. *American Journal of Psychology*, 12(1), 80–130.
6. 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.
7. Bard, E., Shillcock, R., & Altmann, G. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context.

