

Speaking Models: From Levelt's Monolingual to Williams & Hammarberg's Polyglot

Marcia Helena Boëchat Fernandes, Karim Siebeneicher Brito

Universidade Federal do Paraná, Brazil

karimbrito@yahoo.com.br

1. Introduction

Although it is normal for us to produce one or two million word tokens every day, or two to three per second in fluent conversation, we rarely think of the complexity of the speaking process. Researchers, including linguists, psychologists, and neurologists, have attempted to shed some light on the process of speech production and perception, firstly focusing the monolingual speaker, and then advancing to more complex situations, which include bilingual and multilingual speakers.

The present paper begins by providing an overview of speaking models which influenced much of the research in the attempt of understanding the speaking process, both of monolinguals and multilinguals (bilinguals here included). We then present Fernandes-Boëchat's Multilingual Role Model based on her Cognitive Chain-Reaction Theory in Foreign Language Learning, and discuss its relation to other studies in TLA.

We use the notion of L3 here for the language that is currently being acquired, or learned, and L2 for any other language(s) that the speaker has acquired, or learned, after his/her L1. Studies in TLA are mainly based on the fact that L2 and L3 differ substantially, and are highly motivated by these differences as well¹. There is a growing awareness that TLA is not a mere extension of Second Language Acquisition (SLA), and researchers acknowledge that trilingualism demands models of its own, rather than being involved in the scope of those developed in the realm of Second Language Acquisition (Grosjean, 2001; Williams & Hammarberg, 1998).

2. Speaking models

In this section, we will present some of the models which have been referenced to as the most cited and influential. We begin with Levelt's monolingual Model of Language Production, and Green's Inhibitory Control Model, both influencing De Bot's adaptation into a bilingual production model. Grosjean's representation of the language mode continuum follows, as a significant contribution for studies in bilingual speech processing. In the sequence, we refer to some models which account for the multilingual speaker, and his unique traits.

The role and contributions of models are encompassing in their reach because they deal with concepts and their interrelationships. They are developed for the purpose of providing a framework of analysis. According to Rosenblueth and Wiener (1945, p. 316), models are a central necessity of scientific procedure, since abstraction is needed to grasp the part of the universe under consideration. Suárez (1999, p. 168), in the introduction to his chapter, explains that "model building is a pervasive feature of the methodology (or methodologies) employed by scientists to arrive at theoretical representations of real systems, and to manipulate reality." He also states that all the contributors to that book agree that the

¹ See Hufeisen (2000, p. 214) for an account of the several factors that make L3 more complex than the acquisition of a first foreign language.

activity of model building is central to scientific practice, and that “much of scientific practice, perhaps the totality of it, would be impossible without models”.

We understand models as explanation attempts that have overcome their status as hypotheses, because of empirical research data; this does not mean, however, that they do not need to continue to be verified by means of further research.

In psycholinguistic studies, it is highly accepted that words are accessed from the mental lexicon, a huge repository, after activating their associated lexical concepts. The activation depends on the perspective the speaker takes and the choices he makes. In Levelt's (1999, p. 227) words, “working models of word production begin where perspective-taking ends: at the activation of a target concept to be expressed”. Levelt, Roelofs and Meyer (2002, p. 279) suggest that word production emerges from a coupling of two systems, the conceptual and the articulatory motor.

Models which try to explain how speaking takes place followed the first systematic psycholinguistic studies of word production, initiated in the late 1960s (Levelt, 1999). These studies were based on error analyses (Cohen, 1966; Fromkin, 1973), and adopted a chronometric approach to word production, which consists in measuring naming latencies, or the time between a stimulus and a response in naming objects and/or words.²

2.1 Levelt's model of language production

Levelt's (1989) model does not follow the trend of using error analysis as its basis, since he argued that a speaking model should also account for the normal process itself, instead of depending only on infrequent derailments of the process, or errors. His methodology involves basically reaction time research, mainly because it invites the development of real-time process models, thus contributing for the reproduction of the natural process in the laboratory.

Levelt (*op. cit.*) developed a model of human speaking in steps, conceiving speech production as a staged process, in three main components, namely, the *conceptualizer*, the *formulator*, and the *articulator*. It goes from the conceptual/syntactic level to the phonological/articulatory domain, at the beginning of articulation. First, a syntactic word, or lemma, is selected from the mental lexicon, by activating its morphological makeup, its metrical shape (*e.g.* number of syllables, stress position), and its segmental makeup. This involves selecting one appropriate word from among tens of thousands of alternatives, and here the context in which the word is being used plays a crucial role. Activation then spreads through the network in a forward fashion, and nodes are selected following simple rules, the morphological encoder selecting segments and metrical structures, and the phonological encoder selecting the nodes whose links correspond to the phonological syllable positions assigned to the segments. This is finally executed by the articulatory system. Levelt has continually expanded his model to deal with new challenges and data.

At the level of word activation, Levelt argues that “only selected lemmas will become phonologically activated” (2002, p. 312). The author raises the point that speakers make use of a self-monitoring device to control what they are saying, and that this happens not only at the level of their overt speech, but the monitoring is also exerted on their internal speech.

Levelt's model has been adapted by several researchers, in their attempt to make it account for bilingual speakers. We turn to these in the next sections.

² For a more comprehensive explanation of the origin of word production research and its references, see Levelt (1999).

2.2 Green's inhibitory control model

Based on reports of brain-damaged polyglot speakers (Albert & Obler, 1978), Green concludes that, in the human brain, "the subsystems mediating the comprehension and production of language are separable and that different functional systems underlie different languages" (Green, 2000, p. 375). This means that, when brain damage occurs, parts of the speech system can be destroyed or isolated; this would explain why brain-damaged polyglot speakers show the ability of communicating normally in some languages, but seem to have lost the capacity of speaking in others. He presents a model (called the "inhibitory control model") for a bilingual speaker, which is restricted to the comprehension and production of words. Referring to other researchers' work, which make clear that the languages one individual speaks cannot be "deactivated" (Altenberg & Cairns, 1983; Mägiste, 1979; Grosjean, 1982). Instead, he proposes that there are different levels of activation, a language system being *selected* (the one which is controlling speech output), *active* (being conferred some kind of role during the process), or *dormant* (exerting no effects in the speaking process, but still residing in long-term memory). This categorization implies that more than one language can be active at the same time, although just one will be selected for speaking, and this control will depend on the speaker's regulation of the process. Green suggests that the model he outlined can be generalized to account for language control in trilingual or polyglot speakers, as well, and invites for further testing of the model by applying it to these groups of speakers, who, he predicts, should show more problems of control due to more languages involved.

2.3 De Bot's global model of bilingual language production

De Bot (2000) was the first to postulate a bilingual language production model based on Levelt's (1989) model for monolinguals. Poulisse & Bongaerts (1994) later proposed their *Spreading Activation Model*, which was based on De Bot's suggestion. The adaptation De Bot makes of Levelt's model is concerned with the whole speaker, and anything that influences his speech; he thus bears in mind the linguistic, psycholinguistic, and sociolinguistic factors to which the speaker is exposed. He follows Green (2000) in the assumption that the languages a bilingual speaks can be activated to varying degrees, being either selected, active, or dormant.

After introducing and exemplifying Levelt's model, De Bot proposes that a part of the conceptualiser, the formulator and the lexicon are differentiated for the speaker's various languages. Because of individual competence factors, another language that is accessible to him may be activated simultaneously to the selected language, the one the speaker has chosen to speak in. This means that the choice of lemmas, the production of surface structures, and the forming of phonetic plans may happen in parallel in the active language as well as in the selected language, but these planned utterances will not be passed on to the articulator.

The next is not a model, but an alternative solution that has been proposed in a specific point: how bilingual speakers position themselves along a continuum ranging from complete monolingualism to complete bilingualism, according to their interlocutors, and the situation they speak in.

2.4 Grosjean's bilingual language modes

"In the monolingual speech mode, the bilingual deactivates one language (but never totally) and in the bilingual mode, the bilingual speaker chooses a base language, activates the other language and calls on it from time to time in the form of code-switches and borrowings." (Grosjean, 2001, p. 2) The language mode is what he calls "the state of activation of the bilingual's languages and language processing mechanisms, at a given point

in time”(p. 2). It is a concept that has been taken into account by other researchers (Treffers-Daller, 1998; Toribio *et al.*, 2005; Dewaele, 2001), because “it gives a truer reflection of how bilinguals process their two languages, separately or together”, and because it is “invariably present in bilingualism research as an independent, control or confounding variable”(p. 2).

Grosjean (2001) represents the language mode continuum, in which language A, or the base language, is the most active; language B is activated to lesser degrees. This variance in activation of language B determines if the speaker’s use of the language is either closer to one extreme, the monolingual mode, or to the other, the bilingual mode; language B, however, is never as much activated as language A. When interacting with monolingual speakers, bilinguals are believed to be usually in a monolingual mode. In this situation, they are said to deactivate their other language. When the interlocutor knows the other language (B), but would prefer not to use it, the speaker would be in an intermediate position. The speaker’s language B would, then, be only partly activated. Speakers would be in bilingual mode when interacting with other bilinguals who share the same two languages; in this situation they could mix the two languages. Both languages are active, language B being slightly less active than language A, as this is the main language of processing.

The movement along this continuum will differ from one bilingual to the other, regarding the extension they take on it, or the situation when they choose a specific mode. And as the author states that “the bilingual has to decide, usually quite unconsciously, which language to use and how much of the other language is needed” (p. 2), it is believed that the language B does not reach any extreme ends of the continuum, that is, is never completely deactivated, and it never reaches the same level of activation as the base language (A).

Although this is not the norm among TLA studies, some researchers believe that models of bilingual speech can be extended to explain trilingualism; Grosjean (2001) and Roelofs (1998) propose extensions of their own models to encompass more than two languages spoken by the same speaker. Multilingualism researchers, however, tend to point out that models of SLA and bilingualism cannot be used to adequately explain the specific aspects of TLA and multilingualism (see Marx & Hufeisen, 2004). These aspects include not only quantitative differences, as the number of languages an individual speaks, or the amount of words in the mental lexicon, but mainly qualitative ones, like individual learning strategies that have been previously developed, during the process of learning the L2.

So far, some models concerning TLA and multilingualism have been presented, each one focusing on different aspects, and thus complementing each other. The following were presented by Hufeisen (2004:78):

1. Sarah Williams/Björn Hammarberg (1998): *Role-function Model*. A psycholinguistic extension of Levelt’s Model of Language Production, in which the authors emphasize the different roles each language assumes in multilingual speech.
2. Maria Groseva (2000): The *Foreign Language Acquisition Model (FLAM)* regards the L2 acquisition process as underlying the building of hypotheses during third language acquisition.
3. Philip Herdina/Ulrike Jessner (2002): The *Dynamic Model of Multilingualism (DMM)* was developed in the framework of Dynamic Systems Theory, in an attempt to illustrate the dynamic character of the process of multilingual language acquisition. Because of this concern, the authors explain not only acquisition, but also language maintenance and attrition. The model takes into account several factors that underlie multilingual acquisition, and the complexity of their relationship.
4. Britta Hufeisen (2003): Her *Factor Model* emphasizes the distinction between SLA and TLA, advocating that each addition of a further language involves the emergence of several factors, which in turn contribute for a higher level of complexity.
5. Franz-Joseph Meißner (2003): *Multilingual Processing Model*. A constructivist model which considers etymological relations between languages as a key to the referring to

recurrent patterns in the languages in order to understand another foreign language one might have never learned before. When referring to a previously learnt Romance language, for instance, the learner is able to formulate hypotheses about the new Romance language, and thus build a so-called “spontaneous grammar” in this target language.

6. Larissa Aronin/Muiris Ó Laoire (2003): The sociolinguistic *Ecological Model of Multilinguality*. In their contribution, based on Sociolinguistics, the authors differentiate “multilingualism”, which refers to the situation itself, from “multilinguality”, a personal characteristic involving an individual’s store of all his or her interlanguages, as well as every aspect of his or her linguistic identity. The term “ecological” refers to the cultural context in which the authors believe research about multilinguality should happen.

2.5 Williams and Hammarberg’s role-function model

From the TLA or multilingual models presented above, this is the only one which refers exclusively to speech production. The model describes the different roles that the L1 and the L2 occupy in the activation process of L3, and shows they are traceable to different stages of the speaking process, not necessarily involving the same background language. It is based on long-term conversation data collected from Sarah Williams, a native speaker of English who had learned French, Italian and German before starting to learn Swedish, the target language under study. During the whole study, her interlocutor was the other author, Björn Hammarberg. As pointed out by Hammarberg (2001, p. 25), “It (*the corpus*) was compiled without preconceived notions as to what we were going to find. But it gradually became apparent that the third language situation, which was not our initial focus, and especially the roles of L1 and L2 in the acquisitional process, constituted a prominent aspect of Sara William’s handling of Swedish”.

This case study revealed that Williams tended to activate her L1 and L2 knowledge to a considerable degree, and demonstrated a characteristic division of roles between the various background languages in the process. In her case, L1 dominates in various pragmatically functional language shifts, being used to support the interaction or the acquisition of words and other expressions. This was called an “instrumental role”, L1 functioning as an “external instrument language” in the conversations in Sweden. L2, by its turn, occupies a “supplier role” in the learner’s construction of new words in L3, as well as in her attempts to cope with new articulatory patterns in the target language. This role, however, decreases with the increase of L3 proficiency, the target language gradually taking over both instrumental and supplier roles.

One aspect Hammarberg (2001) found remarkable in the results of the study was the strong tendency for just one language to predominate in the role of external supplier, in this case German. He suggests two possible explanations for this: In Sarah William’s case, the German language accumulates more factors (e.g., recency, level of proficiency, L2 status) that condition it for this role; and because it is active on a more regular basis, it is also assigned a more constant role of supplier.

2.6 Fernandes-Boechat’s multilingual role model

This multilingual model is based on the empirical findings of research carried out in France and Brazil since 1997 regarding the role of the preceding foreign language in individual multilingualism (Fernandes-Boechat, 2000; 2006; 2008). There is the strong tendency for just the preceding foreign language to predominate in the role of external supplier during initial target language production, tendency which in fact only occurs when one has achieved at least

intermediate levels of fluency of the preceding foreign language. The model describes the role that the preceding foreign language occupies in the activation process of the target language. Each new foreign language learning experience is linked, involuntarily or unconsciously, by the learner to one's preceding foreign language learning experience in a chain-like domino-effect fashion and as multilingual learners advance from intermediate to higher levels of proficiency in their target language studies, the less they will involuntarily refer back to their preceding foreign language. Unintentional intrusions, therefore, decrease with the increase of target language proficiency.

The model refers to the principles of a cognitive theory, the Cognitive Chain Reaction Theory in Foreign Language Learning (The CCR Theory©)³ in the attempt to explain one of the main reasons for difficulties encountered by highly motivated multilingual learners. Research studies focused on multilingual learners from several countries (*Cultural Diversity*) who converged to a highly qualified center in France to learn French as a foreign language (*Linguistic Unity*) and on Brazilians (*Cultural Unity*) who wanted to learn different foreign languages at Brazilian Universities (*Linguistic Diversity*).

The results of a task performed by multilingual learners, elaborated and administered to measure the learner's performance in the first semesters of target language study, identify the foreign language which learners mostly refer to, involuntarily, during target language speech production. The measurement device utilized to confirm and validate the claims of the *CCR Theory* behind the Multilingual Role Model aimed at observing and identifying the syllable stress produced by the subjects to cognate keywords inserted in a target language text during a recorded reading task. Research findings confirm involuntary preceding foreign language referral during target language production and such referral could, in fact, be one of the main reasons for the difficulties encountered by many multilingual learners and in turn, most probably one of the main reasons for any conspicuous drop-out rate registered among foreign language learners within the first two semesters of target language study.

Some of the results reported by Williams & Hammarberg (1998) corroborate Fernandes-Boechat's findings concerning the stronger influence of the immediately preceding foreign language. These refer specifically to one of the types of non-adapted language switches among the four identified by the authors in a two-year longitudinal case study of Williams' L3 Swedish acquisition. Her L1 was English, and she had three L2s: French, German, and Italian. She studied French and German at university in England, spent one year in France, and took an intensive one-month course in Italian. But after finishing university, she lived in Germany for six years, and went to Sweden directly after that period of time. This makes German the previously learnt language, before Swedish, and also the L2 in which she was most fluent.

The type of language switches identified we referred to was called *Without Identified Pragmatic Purpose*, and shortened to WIPP. The authors believe these correspond to Poulisse & Bongaerts' (1994) 'non-intentional language switches', and in a later work, Hammarberg (2001, p. 38) explains that they do not seem to be caused by a language choice, but rather seem to be lapses in the formulation in L3. The non-intentionality of the productions, though, is not the only aspect of transfer we think strongly relates Williams & Hammarberg's results to Fernandes-Boechat's, but also the fact that 92 per cent of all WIPP switches were German, while only 4 percent were from English, and 4 per cent from other L2s. This strengthens the hypothesis that it is the immediately preceding foreign language which will exert the highest influence on the initial L3 production, and it could also be the case that these intrusions are unintentional. Hammarberg's (2001, p. 38) conclusion is one of surprise that in his study there was "such a strong tendency for *one* language to predominate in the role of external supplier".

³ The CCR Theory was developed in 1997 during Fernandes-Boechat's post-doctorate study leave in a research project at the Centre Universitaire d'Etudes Françaises at Université Stendhal Grenoble III – France.

3. Conclusion

Levelt's monolingual model, as well as De Bot's and other bilingual models have been exhaustively investigated and their propositions evaluated for decades in empirical studies. TLA and multilingual speech models need the same attention from researchers in order to have their assumptions verified, especially in longitudinal studies, like the one developed by Williams and Hammarberg as well as those developed by Fernandes-Boechat. These can lead to great contributions to the field, because they allow the researcher to track change, over a period of time, in specific aspects under study. Hufeisen (2004:146) states that "a research concern in this area is the almost complete lack of empirical studies which attempt to test predictions made by these models".

There are various aspects that need to be investigated further, since so many different factors affect speech, be they its participants or the situation in which it happens. Few studies have focused on pronunciation, for example, and some more attention should be given to subjective factors like psychotypology, introduced by Kellerman (1977), which according to Fernandes-Boechat and Brito (2006), was considered to be the most relevant factor in determining what learners are willing to transfer, and what they eventually transfer from one language to the other.

In this relatively under-explored field, many variables converge to cause cross linguistic influence and interference. Observation of how foreign languages interact with one another during the language learning processes could shed light on how the potential of linguistic knowledge previously acquired can be explored to enhance the learning of other languages.

References

- Albert, M. L., & Obler, L. K. (1978). *The bilingual brain: Neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.
- Altenberg, E. P., & Cairns, H. S. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 22, 174-188.
- Cohen, A. (1966). Errors of speech and their implications for understanding the strategy of language users. *Zeitschrift für Phonetik*, 21, 177-181.
- De Bot, K. (2000). A bilingual production model: Levelt's 'speaking' model adapted. In L. Wei (Ed.), *The bilingualism reader* (pp. 420-442). London; New York: Routledge. (Reprinted from *Applied Linguistics*, 13, 1992, 1-24.)
- Dewaele, J.-M. (2001). Activation or inhibition? The interaction of L1, L2 and L3 on the Language Mode Continuum. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 69-89). Clevedon: Multilingual Matters.
- Fernandes-Boechat, M. H. (2000). The cognitive chain-reaction theory in foreign language learning. In *AILA (Association Internationale de Linguistique Appliquée) '99*, CD-ROM. Tokyo, Japan.
- Fernandes-Boechat, M. H. (2006). O aprendiz multilíngüe e a teoria cognitiva de reação-em-cadeia. In A. M. Karwoski, & V. de F. C. V. Boni (Eds.), *Tendências contemporâneas no ensino de línguas* (pp. 33-43). União da Vitória: Kaygangue.
- Fernandes-Boechat, M. H. (2007). The CCR Theory: A cognitive strategy research proposal for individual multilingualism. *Revista Luminária*, 8(1). FAFIUV: União da Vitória-PR. Retrieved 24 January 2008, from <http://www.ieps.org.br/luminaria.pdf>.
- Fernandes-Boechat, M. H. & Brito, K. S. (2006). A psico-tipologia e suas implicações no aprendizado de línguas estrangeiras. *Revista Horizontes de Linguística Aplicada*, 5 (2), 25-36.
- Fromkin, V. A. (1973). *Speech errors as linguistic evidence*. The Hague-Paris: Mouton de Gruyter.

- Green, D. W. (2000). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. In L. Wei (Ed.), *The bilingualism reader* (pp. 374-385). London; New York: Routledge.
- Grosjean, F. (1982) *Life with two languages*. Cambridge: Cambridge University Press.
- Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1-22). Oxford: Blackwell.
- Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 21-41). Clevedon: Multilingual Matters.
- Hufeisen, B. (2000). A European perspective: Tertiary languages with a focus on German as L3. In J. W. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 209-229). New Jersey, USA: Lawrence Erlbaum Associates.
- Hufeisen, B. (2004). „Das haben wir doch immer schon so gemacht!“ oder ein Paradigmenwechsel in der Spracherwerbsforschung? In K.-R. Bausch, F. G. Königs, & H.-J. Krumm (Eds.), *Mehrsprachigkeit im Fokus: Arbeitspapiere der 24. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts* (pp. 77-87). Tübingen: Gunter Narr Verlag.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223-232.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (2002). A theory of lexical access in speech production. In G. T. M. Altmann (Ed.), *Psycholinguistics: Critical concepts in psychology* (pp. 278-377). London: Routledge, v. 5. (Reprinted from *Behavioral and Brain Sciences*, 22, 1999, 1-38.)
- Mägiste, E. (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior*, 18, 79-89.
- Marx, N., & Hufeisen, B. (2004). A critical overview on third language acquisition and multilingualism published in the German language. *International Journal of Multilingualism*, 1 (2), 141-154.
- Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36-57.
- Roelofs, A. (1998). Lemma selection without inhibition of language in bilingual speakers. *Bilingualism: Language and Cognition*, 1, 94-95.
- Rosenblueth, A., & Wiener, N. (1945). The role of models in science. *Philosophy of Science*, 12(4), 316-321.
- Suárez, M. (1999). The role of models in the application of scientific theories: Epistemological implications. In M. S. Morgan, & M. Morrison (Eds.), *Models as mediators: Perspectives in natural and social science* (pp. 168-195). Cambridge: Cambridge University Press.
- Toribio, A. J. et al. (2005). Perseverative phonetic effects in bilingual code-switching. In R. S. Gess, & E. J. Rubin (Eds.), *Theoretical and experimental approaches to romance linguistics: Selected papers from the 34th Linguistics Symposium on Romance Languages (LSRL)* (pp. 291-306). Salt Lake City: John Benjamins.
- Treffers-Daller, J. (1998). Variability in code-switching styles: Turkish-German code-switching patterns. In R. Jacobson (Ed.), *Codeswitching worldwide* (pp.177-200). Berlin; New York: Mouton de Gruyter.
- Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics*, 19, 295-333.