John Benjamins Publishing Company

This is a contribution from *Review of Cognitive Linguistics 12:1*
© 2014. John Benjamins Publishing Company

This electronic file may not be altered in any way.
The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.
Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).
Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Fuzzy sets and Prototype Theory

Representational model of cognitive community structures based on lexical availability trials*

Antonio M. Ávila-Muñoz and José M. Sánchez-Sáez
University of Málaga

Prototype Theory offers one of the most accepted models for semantic memory organization. Lexical availability trials provide investigators with a faster and easier means of observing this cognitive organization, since lists of available lexicon are generated from associations relating some lexical elements with others. The experiments with lexical availability are able to activate one of the best-known lexical production mechanisms within experimental psychology: semantic category fluency. In this work we propose an appropriate means to reconstruct the community cognitive organization. This shared metastructure constitutes the concept of shared field of experience used as the base for availability trials. The key notion is the prototypicality of common vocabulary as the base for the construction of community models. To obtain a representation of these prototypes we use the mathematical framework of fuzzy sets.

Keywords: Prototype Theory, lexical availability, fuzzy sets

1. Introduction

For the scientist, the structural organization of an individual's lexicon is a type of black box, the content of which is not directly observable.[1] If we wish to create models that represent this opaque structure, we must carry out experiments that

* The work reported here is part of the research undertaken within the framework of DGICyT Research Project on the Sociolinguistic Patterns of Málaga Spanish (Patrones Sociolingüísticos del Español de Málaga, PASOS-Ma; FFI2011-29189-C05-01), and has benefited from the grants that fund it.

1. The classic analogy of the black box (Haber-Schaim, Cross, Abegg, Dodge, & Walter, 1979) is a helpful comparison for understanding the importance of the creation of scientific models

Review of Cognitive Linguistics 12:1 (2014), **133–159.** DOI 10.1075/rcl.12.1.05avi
ISSN 1877–9751 / E-ISSN 1877–976X © John Benjamins Publishing Company

tell us about its form. One of the most frequently used is the available lexicon, which inserts a point into the lexical network. These trials enable us to study the lineal spread through the lexicon from the insertion point. However, this does not indicate the exact form of the lexicon itself which is usually dynamic and organized in multiple levels and directions that are difficult to control.

The preparation of lists of lexical availability has always had an important place in the context of lexical statistical research, especially in the Hispanic world. The reasons for this deep-rooted tradition lie in the multiple applications that researchers usually find in the lists of availability. Disciplines such as sociolinguistics, ethnolinguistics, dialectology, psycholinguistics or pedagogy, among others, have used studies on availability in many and effective ways.

In spite of their deep-rooted tradition, studies on lexical availability lack a general theory which allows the expression of a theoretical framework and answers all the questions essential to devising this type of list. The root of the problem lies in the early works on lexical availability which contained no detailed theoretical explanation of the lexical basis of the study, nor of how the mental lexicon was organized; what the list of a speaker's lexical availability represented or what cognitive, individual or collective structure it reproduced. Answers to these key questions would have helped to draw conclusions on the mental structure of lexicon and the use of the lexicon in speech communities.

In this paper, we use the essential nature of the trials on available lexicon to gain a better understanding of the mental lexicon of speakers. We shall start from this hypothesis: when an individual takes part in an experiment on lexical availability, mechanisms of association reproducing cognitive models typical of the prototype theory are triggered in his or her mind. The analysis of lexical networks of a group of effectively-linked individuals shows similar forms because of their socio-cultural determinants. This shared metastructure, or *shared field of experience*, will allow us to construct a model capable of representing the shared lexical spectrum. We propose to study this representation using the theory of fuzzy sets.

We will apply the model chosen to determine the lexical production mechanisms observed in a sample of speakers (N = 72) who made up the lists of available lexicon in the city of Málaga in southern Spain, between the years of 2002 and 2005. This sample is described in Section 4.1.

Two well-established and widely accepted lines of reasoning support the theoretical framework of this work:

---

that explain reality through experiments based on hypotheses over a subject that is not directly accessible.

© 2014. John Benjamins Publishing Company
All rights reserved

1. Firstly, the Research Project on the Spanish of Spain and the Americas (Proyecto PRESEEA, headed by F. Moreno Fernández, Cervantes Institute/ UAH). This project aims at studying sociolinguistic variation in a broad selection of urban communities in Spain and the Americas. In the framework of the PRESEEA research for Málaga (PASOS-MA), our current study takes advantage of an adapted version of the sociological and network questionnaire used in the community study.
2. Secondly, the so called 'Pan Hispanic Project for the Study of Lexical Availability' (Proyecto Panhispánico de Disponibilidad Léxica, PPHDL). This macro-project is aimed at obtaining large amounts of data from students in the final year of High School, and analyzes their available vocabulary following basic common criteria. The results are compared in order to build up a 'Diccionario Panhispánico de Disponibilidad Léxica' (Pan Hispanic Dictionary of Lexical Availability).

A brief revision of the theoretical framework on lexical availability is shown in Section 2. In Section 3 we propose an alternative model for this issue. Section 4 focuses on the proposed process and variations to build up the availability index. Conclusions are presented in Section 5.

2. Lexical availability

2.1 Concept

The first studies into available lexicon appeared in the mid 20th century, in a linguistics program sponsored by UNESCO. It was aimed at helping the native population in the French colonies and immigrants to France to learn French. It was anticipated that a command of a lexically and grammatically simplified "basic language" would facilitate these citizens' social integration. This education program was centered essentially on vocabulary, so the importance of the tokens selected was stressed from the outset.

Available lexicons are considered a reflection of the lexical resources speakers use in specific communicative situations. This type of list operates using the concept of *frequent situation*. The significance of the availability lexicon lies in evidence that certain words used frequently in one specific language are closely connected to the appearance or otherwise of certain topics throughout communicative interactions. Words such as *stage, proscenium, actor*, *stalls* usually appear only in communicative exchanges where matters related to the theater are being discussed.

© 2014. John Benjamins Publishing Company
All rights reserved

2.2 Tradition

The work serving as a basis, model and departure point, marking the guidelines for later studies on available lexicon, was *L'élaboration du français fondamental*, written by Gougenheim, Michéa, Rivenc, and Sauvageot (1956). Experts carrying out research in lexical availability in different languages immediately grasped the possibilities of these types of work. In 1969, Dimitrijêvic (1969) designed a study on the available lexicon in Scotland. Shortly afterwards, in 1971, Mackey and his associates (Mackey, 1971) compared Gougenheim's conclusions with those they had obtained analyzing the available lexicon in the spoken French in Canada. Logically, to carry out a coherent comparison, Mackey and his associates assumed the methodology previously outlined by the European pioneers. The result was the publication of two volumes establishing an intralinguistic comparison covering areas of intraculturality. This research was carried out in the *Centre Internacional de Recherches sur le Bilinguisme* de L'Université Laval de Québec.

Works written and advanced by Humberto López-Morales have helped to consolidate the tradition of lexical availability studies in the Hispanic sphere (López-Morales, 1973, 1978, 1979 and above all, 1999). To them we owe the establishment of the Panhispanic Lexical Availability Project (*Proyecto Panhispánico de estudio sobre la Disponibilidad Léxica*, PPHDL). Its aim is to obtain a dictionary of Spanish available lexicon from lists of lexical availability in different Spanish-speaking areas in the world.

The importance of this coordinated macro project lies, above all, in its proposal to equate and unify the criteria of the methodological design. This permits interchange of data and development of comparative studies of results from different local projects.

The various groups taking part in the macro project adopt methodology guidelines summarized as follows:

1. Participants: pre-university students. This avoids any specific or "technical" contamination from a particular professional field and focuses the study on subjects presumed sufficiently "mature" from a lexical point of view (Benítez, 1992, p. 73; López-Morales, 1999, p. 28).
2. Material is obtained using associative trials from sixteen centers of interest.[2] The informant must provide the lexical elements he or she considers related.

2. The term 'center of interest' has been used since the beginning of availability studies, but a consistent definition of its nature and meaning is lacking. In reality the concept of 'center of interest' is equivalent to 'prototype' in the classic bibliography. For this reason, we will use the term 'center of interest' when classic availability studies are refereed and a new terminology will be used for the work proposed in this paper.

© 2014. John Benjamins Publishing Company
All rights reserved

The centers of interest proposed are those used in the pioneering works of Gougenheim: 01. The Human Body; 02. Clothing; 03. Parts of a house (not including furniture); 04. House Furniture; 05. Food; 06. Objects on a table laid for a meal; 07. The Kitchen and Kitchen Utensils; 08. The School: Furniture and Equipment; 09. Heating, Lighting and Means of airing a Premises; 10. The City; 11. The Countryside; 12. Means of Transport; 13. Farm or Garden Work; 14. Animals; 15. Games and Entertainment; 16. Trades and Professions. However, there have always been discrepancies as to the number and class of appropriate centers of interest for the task (Butrón, 1987). Perhaps the most critical and realistic approach can be seen in Hernández-Muñoz (2006, pp. 127–188) who bases his exposition on: (1) not all the centers of interest respond to the traditional concept of natural categories organized around a prototype (internal organisation of center of interest); (2) centers of interest can have an inclusive character: the tokens are elements on a basic or subordinate level, or they can have a relational character: the tokens do not establish a direct belonging relationship (inclusivity levels). As there may be times when this identification becomes difficult, subjects may develop response strategies that manipulate the results. This leads Hernández-Muñoz (2006, pp. 185–186) to declare that neither in quality nor in quantity does unity exist between the center of interest selected for the studies on available lexicon and, most important of all, that the token availability may be due to different factors. From this it is easy to conclude that availability is not a permanent quality of words (it is probably not even in their nature), and that it depends to a great measure on the categorical framework (Hernández-Muñoz, 2006, p. 186).

3. The lists are open inasmuch as there is no limit to the number of words submitted, although the informant must write down as many as possible related to the subject proposed.
4. A time limit is the only condition imposed. Individuals must not take more than two minutes to write their lists on each center of interest. Presumably, the first words that come to mind are the most available and will be at the top of the lists.
5. The content of the word lists are edited using previously-agreed lemmatization criteria. The statistical analysis of the databases provides us with the final lists of lexical availability. To obtain these results, researchers use free distribution software developed by the University of Alcalá, Madrid, Spain, and approved by the Cervantes Institute: *LexiDisp* is an application for

© 2014. John Benjamins Publishing Company
All rights reserved

Windows© able to calculate the lexical availability using a mathematical formula developed by López-Chávez and Strassburger-Frías (1987).[3]

The social and personal characteristics of the subjects at our disposal make it possible to establish specific sociolinguistic correlations. These will later be compared and contrasted with those obtained in other studies, thus greatly enriching the initial possibilities.

2.3 Cognitive foundations of lexical availability

Leaving aside the controversy over the classical dichotomies in the sphere of Psycholinguistics, there does appear to be consensus on the existence of two levels of representation that contain the information needed to produce the terms associated with a semantic category. One is the mental lexicon. This concept is constantly present in the precarious explanations the majority of lexical availability researchers used to draft an incipient theory on the psychological foundations of their lists. The other is the semantic memory where the first selection of lexical material involved is apparently produced.

The concept of *mental lexicon* is generally understood to be the storehouse of word forms (Levelt, Roelofs, & Meyer, 1999). The concept of *semantic memory* is associated with the representation of the meaning of words and could be understood to be a huge cognitive storehouse (Sartori, Coltheart, Miozzo, & Job, 1994, p. 538). A third concept, that of *semantic priming* (Troyer, Moscovitch, & Wincour, 1997), lies between these two essential elements. It refers to an accessibility process in individual responses that is found in the implicit memory and which manifests itself in the available lexicon.[4] The function of semantic

3. Both the program and the operating instructions can be found at the following link: <http://www.linguas.net/Proyectos/LexiDisp/tabid/73/Default.aspx> [last consulted: March, 16, 2013]. Furthermore, the PPHDL has an electronic meeting place where anyone researching or interested in lexical availability can find the latest new information: <http://www.dispolex.com/> [last consulted: March, 16, 2013]. Dispolex is more than just a mere means of contact, because thanks to its interface, a registered user can incorporate their material in a general data base bank. According to its creators (Bartol, Hernández Muñoz, Salamanca University, Spain) this contributes to the making of an extensive *Panhispanic repository* constituted in such a way as to adjust to the peculiarities of specific research. Furthermore, the registered researcher has access, at no cost, to the tools necessary for making the most common calculations of lexical availability: *availability index, frequency, percentage of occurrence, number of words, number of tokens, informants' averages, cohesion index and project comparisons*.

4. According to Holyoack (2002, p. 37), for example, if the word *saffron* has recently been studied, it is much more likely that the person will generate *saffron* when asked to enumerate flowers even though he does not remember explicitly having studied the word.

© 2014. John Benjamins Publishing Company
All rights reserved

memory is widely accepted by the scientific community as responsible for coding all perceptive and conceptual information, reflecting an individual's implicit vision of the world. Nevertheless, its organization and composition are the center of a heated debate which we shall disregard for the present.[5]

Adapting Levelt'sand Caramazza's theories and models of lexical production to the production of available lexicon (Caramazza, 1997; Levelt, 1999, 2001; Levelt, Roelofs, & Meyer, 1999) leads us to think that the mental processes involved in generating lists of availability occur in the following way (Hernández-Muñoz, 2006, pp. 77 and ff.): when an individual is prompted by a prototype, the semantic memory corresponding to the lexical concept and other associated concepts (*skirt*, *trouser*, *shirt*, etcetera) is set in motion (for example, for the 'Clothing' prototype, center of interest *clothing*). Subsequently, the *lemmas* corresponding to each of these semantic representations are triggered at the mental lexicon level. *Lemmas* are defined as functional units that mediate between concepts and forms and moreover contain the syntactic information of the words. Since one concept may activate different *lemmas* it is assumed that each individual will choose the one best suited to the context at a given moment, determined by multiple parameters. Although subjects do not have to incorporate syntactic information in the availability trials, it appears that this is triggered unconsciously, even though in these types of trial there is no need to set the whole syntactic process in motion. Only the grammatical features are translated as soon as we write or utter a chain of speech (Hernández-Muñoz, 2006, p. 80). In the next phase, depending on whether the message is in written or oral form, the triggering of these *lemmas* spreads towards the phonological or orthographical forms where the final selection process that culminates in the realization of the most available word for each subject takes place. In the process of writing down lists, the level of phonological representation step will not be necessary since it is possible that the semantic level is directly connected to the graphemic 'storehouse' of the language.[6]

---

5. Broadly speaking, we must remember that the most important models of semantic memory organization are the feature theory (Chomsky, 1965; Johnson-Laird, 1983; Katz & Fodor, 1963), semantic networks (Collins & Loftus, 1975; Collins & Quilian, 1969), prototype theory (Lakoff, 1987; Rosch, 1978; Wittgenstein, 1953) and meaning postulates (Carnap, 1952). Of these four representational models of semantic memory, experts in available lexicon have paid special attention to the theory of semantic networks (Galloso-Camacho, 2003; Gómez-Devís, 2003; López-Morales, 1999; Urrutia, 2001) and to the prototype theory (Hernández-Muñoz, 2006; Romero-Rubilar, 2000).

6. Other schools of thought (Ellis, 1982) consider that the phonological level mediates between the semantic and the graphemic. A theory that conciliates both these positions can be seen in Cuetos (1991, p. 38), which maintains that in productive writing both paths are triggered together.

© 2014. John Benjamins Publishing Company
All rights reserved

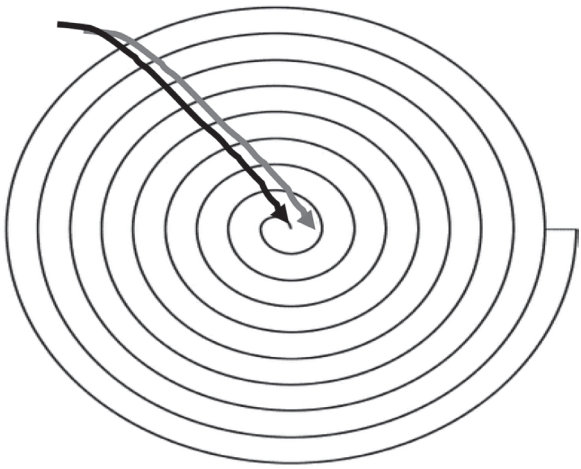


Figure 1. Gateway to lexical network: core tokens

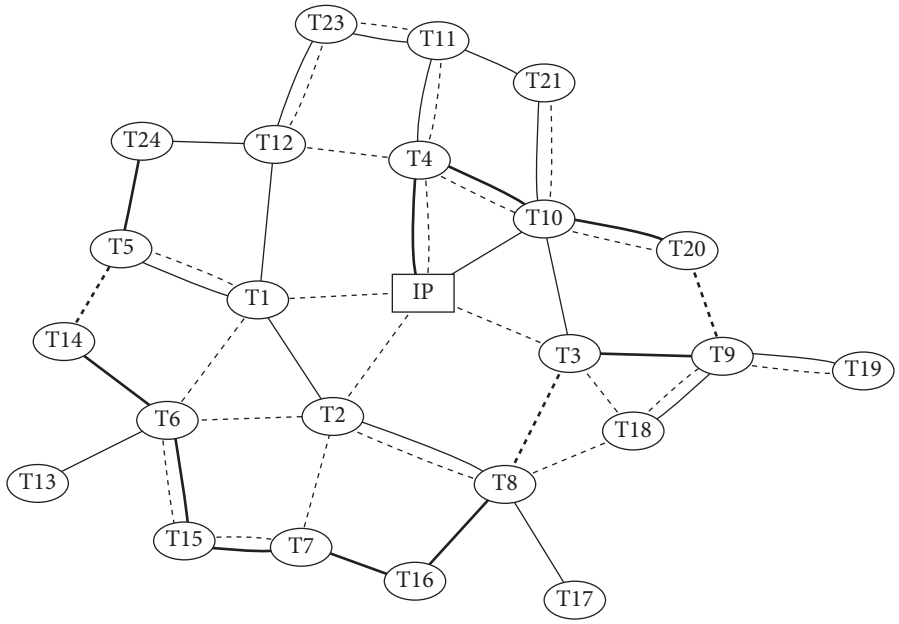


Figure 2. Lexical selection in a multi-connected network

Based on our own empirical observations, we have found out that when an individual takes part in an experiment on lexical availability, they access their lexical network through the prototype. Subsequently, they move through the associated lexical network which will take them progressively further from the

© 2014. John Benjamins Publishing Company
All rights reserved

prototypical nucleus. If we observe this process as a panorama, we could assign it the geometric form of a spiral. At the center, we will find the most central or prototypical tokens the subjects choose when prompted by the prototype. As the trial evolves, the individuals travel this spiral in such a way that, as subsequent terms appear, they gradually lose this condition of being central or core (Figure 1).

However, when we observe what really happens during the individual selection process of terms more closely, we can see that this spiral contains details resembling a multi-connected network on different levels. The Insertion Point (IP) is at the center, and a path travels through the network following certain links that continually move away from the IP (Figure 2). As we have actually observed this continuous motion as a general scheme in every case studied in our research, we feel confident that this behavior is customary in this kind of lexical selection process.

In the example shown, on receiving the IP the individual traces a selection path starting at T4 and thence triggering T10, T20, T9, T3, T8, T16, T7, T15, T6, T14, T5 and T24 (represented by the thick line). Figure 2 shows that all the terms are linked in the mind of the individual by different processes of association represented by different types of lines.[7] For example, solid edges could correspond to semantic similarity or closeness and dashed lines to phonetic resemblance.

This interpretation allows us to reinterpret the trials under the framework of cognitive studies. This approach is innovative and original, as it aims at providing a new interpretation of this kind of trials.

2.4 Lexical availability index

The availability index is, broadly speaking, a quantitative referent that attempts to relate frequency and order criteria to obtain results compatible with the concept expressed in Section 2.1. The mathematical calculation on which it is based is aimed at weighting token frequency realization in a center of interest with its position in the various lists. The starting point is a matrix similar to that shown in Table 1, where v^1, v^2… are the *lemmas* produced by the speaker in a given center of interest.

Subjects 1, 2 and 3 made lists of different lengths. The tokens included in individual lists did not always match, and when they did, the tokens were not

7. The characteristic words of a center of interest are connected by semantic content or field value (first associative axis). However, within these lexical fields, speakers resort to specific associative options from the strictly lexical to semantic, morphological or phonological procedures. For a review of the different mechanisms of word association detected in availability trials see Galloso-Camacho (2003, pp. 133–163).

© 2014. John Benjamins Publishing Company
All rights reserved

Table 1. Matrix of token occurrence by subject. Source: adapted from Butrón (1987, pp. 23–35)

Subject 1	Subject 2	Subject 3
v^1	v^2	v^1
v^3	v^4	v^2
v^2	v^1	v^4
v^5	v^3	
	v^5	

Table 2. Token position vector. Source: adapted from Butrón (1987, pp. 23–35)

v^1	20100
v^2	11100
v^3	01010
v^4	01100
v^5	00010

always in the same position. Using the previous matrix it is possible to construct a position vector for each of the tokens, as shown in Table 2.

From this it can be concluded that token 1 (v^1) was noted down by the population group under study twice in first position, never in second, once in third and never in fourth or fifth. Token 2 (v^2) was noted down once in the first three positions but never in the last two, and so on.

This process shows that only the token position and the frequency of occurrence in that position are relevant to determine the lexical availability index. In fact, since the works of Lorán (1983) and Lorán and López-Morales (1983), strategies for weighted frequency have been designed that included token position as a determining factor of the availability index. After several revisions, López-Chávez and Strassburger-Frías (1987) proposed a definitive formulation for this calculation. The basis was a vector matrix in which the following are included:

a. the maximum number of positions reached in the center of interest under analysis: n
b. the positions where the units were noted: $i = 1,\ldots,n$
c. the absolute frequency obtained by the unit (p) in each position (i) on the lists: f_{pi}
d. the number of speakers that reach the position i: N_i
e. the number of individuals taking part in the study: $n1$
f. a control parameter: C

The final expression of the proposed formula is thus established as:

$$d(p) = \sum_{i=1}^{n} \exp\left(C - \frac{i-1}{n1}\right)\frac{f_{pi}}{N_i}. \quad (1)$$

The application of the exponential function with a complex exponent is responsible for a true weight of the position and is what conveniently permits combining

© 2014. John Benjamins Publishing Company
All rights reserved

the frequency and position of each unit because, whatever the number of participants, it offers a variable weight between 0.1 and 1 of the extension of each individual list and each lexical element frequency of occurrence.

In spite of the fact that this formula represents a major progression, a certain number of cases – mostly theoretical – have been raised for which the behavior of the formula is not adequate. Particularly, the formula tends to cause some trouble whenever lists of different length are considered (López-Chávez & Strassburger-Frías, 1991).

3. Lexical availability and prototypicality

3.1 Fundamentals

We shall now formulate the first main hypotheses that lead to the construction of an abstract model of prototypical community representation. Several of these have been sufficiently tested in other studies and are taken here as the basis for developing this research.

According to the psycholinguistic tradition, the *mental lexicon* is similar to a mental network that contains some information regarding a word's meaning, its pronunciation or its syntactic features. However, as the active nature of the mental lexicon keeps it constantly updated, it is hard to identify how words are interlinked and which is the easiest way to access it. Discussion of this theoretical issue is beyond the scope of this paper, therefore we have chosen the simplest and most reliable model.

Experiments usually begin with the proposal of a prototype (let us say, 'Food') that is able to cause the production of an individual list of related *lemmas* (a, b, c…). By this procedure the mental lexicon arranges structures for itself around the proposed prototype. That is what we refer to as *field of experience*. The proposed prototype provides a stimulus or access point that is used by the individual to enter his or her network of lexical elements.

On the one hand, the structure of the field of experience is personal and inherent for every speaker, and should not be strictly extrapolated to anyone else. In the same way as the mental lexicon, the structure of the field of experience changes for every speaker and over time. Therefore, it is very likely that the same speaker may generate different lists for the same prototype at different times.

On the other hand, speakers tend to produce closely parallel lists. This produces the impression that the networks of lexical elements from different speakers would have the same words adjacent to the proposed prototype, even though the web of links in each speaker's network is unlikely to be the same. This shared

© 2014. John Benjamins Publishing Company
All rights reserved

metastructure is what we call *shared field of experience*, and it may be identified as our main subject of study. This concept would correspond to the main meaning of the classic term '*center of interest*'. Because of the tradition and fuzzy concepts involved in the use of the term '*center of interest*', often indistinguishable from *prototype* or the proposed concept of *shared field of experience*, we decided to coin new terms that respond to the model exposed in this work.

However, a network of links between words does not seem to be the most adequate model to represent the structure of a shared field of experience, since this sort of lexical network is supposed to belong to the inner self of every speaker. When dealing with lexical availability, we need a model involving the distance of a word to the prototype provided by the stimulus. This proximity, interpreted as an availability index for every word in the sample, is what is called *linguistic spectrum*, and it is the proposed solution to the problem of deciding when a word belongs to the core of a shared field of experience (i.e., if it is equivalent to the prototype). Thus it is necessary to choose a representation model capable of expressing the compatibility with the concept represented by a given set of words, and providing different levels of compatibility. The concept 'compatibility' is closely related to the concept of 'accessibility' (Ávila-Muñoz & Sánchez-Sáez, 2010, pp. 35–82). In fact, quantification of accessibility for every lexical entry in the shared field of experience will correspond with its availability, once the information gathered from all the speakers of the sample is incorporated.

Compatibility between words and their degree of accessibility determine one basic aspect of the representation of the lexical structure of a shared field of experience. Thus, the more available a word is, the greater and quicker its accessibility will be (that is to say, a greater availability index).

3.2 Availability trials

Our work is based on a theoretical model intended to build up the foundations for more objective and realistic tests of lexical availability (Ávila-Muñoz & Sánchez-Sáez, 2010, 2011; Villena-Ponsoda, Ávila-Muñoz, & Sánchez-Sáez, 2012). In trials to obtain available lexicon for a shared field of experience: (1) access is gained to a lexical element close to the proposed prototype; (2) as the words are produced, a progressive distancing from the main concept, via related elements, takes place; (3) when the speaker considers that he or she has moved too far from that prototype, he or she tends to choose a fresh word close to the prototype. This re-entry process means that each individual sample will not be strictly determinant of the individual's lexicon structure; (4) the lexical capacity of each speaker (size and originality of their vocabulary) and the speed of access are determinants in the

© 2014. John Benjamins Publishing Company
All rights reserved

process, although in the classic availability trials the speed of access to the lexicon is not controlled. Furthermore, the speed of access may be conditioned by elements extrinsic to the trial that could have been interposing while the experiment was taking place; (5) the order of occurrence of the words, although non-determinant, is relevant as regards its accessibility.

3.3 Lexicon structure

Associations within each speaker's lexical network are unique, but people tend to associate the same words to the proposed prototype. Therefore, a shared field of experience is not a network by itself, but a set of availability degrees of a certain word with reference to the prototype.

The lexical structure of a shared field of experience is considered to consist of: (1) the nucleus or set of basic terms accessible to all speakers; (2) the periphery or set of terms not accessible to all speakers; (3) a gradual rather than abrupt progression from nucleus to periphery, where we observe a gradation rather than a dichotomist classification. The belonging of lexical elements in the shared field of experience can be interpreted as the *compatibility* or *affinity* of each lexical element with the prototype represented by the shared field of experience.

3.4 Fuzzy sets as representation of the availability index

A main assumption of this work is that each shared field of experience emerges and revolves around a prototype. This prototype has already been produced from the source-concept of the shared field of experience itself. As a matter of fact, our proposal derives from the prototype theory (Lakoff, 1987; Rosch, 1978; Wittgenstein, 1953) and considers that the availability index is an excellent indicator of the degree of prototypicality that each token has within each of the shared fields of experience. It is advisable not to associate the availability index with the nature of the tokens because, to a great extent, this index is linked to the collective conceptual categorization. According to our premise, the availability of a token in a shared field of experiencecorresponds essentially to the concept of *accessibility*.

The availability index taken from the individual lists may allow us to calculate quantitatively the structure of the accessibility to the lexicon for a population group in a given shared field of experience. Quantification of this accessibility is the measure of the concept of availability of each term in that shared field of experience, once the information provided by all the individuals in the population group has been incorporated.

© 2014. John Benjamins Publishing Company
All rights reserved

The association between the words and the accessibility displayed by each of these determines how the lexicon structure in a shared field of experience is represented. In this representation, the tokens closest to the prototype of the thematic nucleus will show a higher degree of accessibility (that is, a higher availability index).

In our approach we need to establish the degree of *compatibility* between, on the one hand, each of the tokens contributed by the speakers and on the other, the concept represented by each of the shared field of experience proposed for the community under scrutiny. This quantitative compatibility is represented through the mathematical concept of *fuzzy set* (Zadeh, 1965, pp. 338–353; Zimmermann, 2001).

Broadly speaking, the fuzzy sets are a generalization of the set theory in which, instead of considering their belonging, the compatibility of the elements with the concept represented by the set is considered. In this way it is possible to establish different levels of compatibility between an element and the set against which this compatibility is measured, instead of the belonging-not belonging dichotomy typical of the classic set theory. This concept of *compatibility* corresponds, in our case, with the concept of *accessibility*.

This approach allows us to take advantage of a previously established and contrasted mathematical framework that enabled us to use the tools available in that field. This helps us resolve other problems originating from those already raised with the basic organisation of the terms. Until now, these types of problem have appeared incidentally to occupy a place of preference among available lexicon research projects (Ávila-Muñoz & Sánchez-Sáez, 2010).

For example, one of the tools that bring the theory of fuzzy sets within our reach is determining the characteristic compatibility of the fuzzy set by using either FEV – Fuzzy Expected Value – or its variation, WFEV – Weighted Fuzzy Expected Value – (Friedman, Schneider, & Kandel, 1989); i.e., which compatibility value (availability index) is considered to be a lower bound for a word to belong to the shared field of experience. By using this we can establish, on the one hand, a limit of characterization of belonging values and on the other, parameters for identifying "very characteristic" or "hardly characteristic" elements. In fact, it is a case of proposing an objective cut-off mark in the higher and lower levels of the fuzzy set that does not depend directly on the subjective perception of the researcher. This mark will be given in direct relation to the degree of compatibility of each element with an assessment of the set of elements selected, this last factor enabling us to parameterize the differentiating process.

To sum up, understanding the lexical availability index through compatibility measures from Fuzzy Sets Theory provides a wide framework filled with tools able to deal with the data conclusions obtained as well as any additional information.

© 2014. John Benjamins Publishing Company
All rights reserved

4. Determining the shared lexical-cognitive structures (lexical spectrum)

In order to state the lexical spectrum of a particular shared field of experience, a constructive approximation has been chosen. The result of this process can be interpreted as a quantitative evaluation of the compatibility of each token with the prototype proposed for this shared field of experience. Determining this information means solving the problem of obtaining the availability of each element. By considering the availability of an element as a degree of compatibility, the problem is translated into the evaluation of these degrees of belonging.

4.1 Sample

The data serving as basis for this work were obtained and analyzed throughout the Research Project on Social Conditioning of the Málaga Available Lexicon (*Proyecto de Investigación sobre el Condicionamiento Social del Léxico Disponible en la ciudad de Málaga*, CONSOLEX).[8]

The general aim of the research project was to produce a dictionary of availability and a parallel study of social variation in speakers' available lexicon in the city of Málaga. We used a uniformly allocated sample of speakers, stratified by age, sex and education level (n = 72), from a total population census of 557,770 inhabitants in 2005, giving a sample/population ratio of 1/7746, which goes far beyond the accepted representativeness among the main sociolinguistics studies. Furthermore, the proportion of the sample used in this project improves the established minimum limit established in the general framework for our research (Proyecto PRESEEA: 1/25.000).[9] Table 3 shows the sample distribution.[10]

On the one hand, PRESEEA-Project methodological principles suggest using sampling by quota instead of random or probabilistic sampling, since this method of selecting informants makes for easier statistical comparison in the same sample and between different samples. However, it is interesting to bear in mind that the

---

8. Excellence Project (HUM-315), Andalusian Autonomous Government, Ministry of Science and Innovation, Spain.

9. Criteria used were taken from those laid down in the framework of the PRESEEA Project. The theoretical and methodological principles of this Panhispanic project can be consulted in the following link: <http://www.linguas.net/portalpreseea/Metodolog%C3%ADa/tabid/474/Default.aspx> [last consulted: March, 26, 2013].

10. In point of fact the objective was in itself a novelty since, as mentioned above, lexical availability studies participating in the PPHDL macro project only study lexical availability in pre-university students.

© 2014. John Benjamins Publishing Company
All rights reserved

Table 3. Distribution of sample of speakers in the study on social variation in available lexicon in Málaga (CONSOLEX Project).

Man					Woman				
Education	**0**	**1**	**2**	**Total**	**Education**	**0**	**1**	**2**	**Total**
Age					Age				
20–34	4	4	4	12	20–34	4	4	4	12
35–54	4	4	4	12	35–54	4	4	4	12
> 55	4	4	4	12	> 55	4	4	4	12
Total	12	12	12	36	Total	12	12	12	36

main objective of PRESEEA is urban speech variation in Spain and in the Spanish-speaking countries of America. For that purpose a set of unified theoretical and methodological criteria have been generated, in order to have access to sociolinguistic patterns of speech used in most of the largest cities and towns in both Spain and Latin America.

On the other hand, data to be analyzed in our word availability project were compiled according to the normal procedure from the macro-project PPHDL (*supra*, 2.2). In our case, keeping the customary previously-used sixteen fields of shared experience, we added four other specific centers or prototypes: 17. Money and Economy, 18. Internet and Computers, 19. Colors, and 20. The Sea.

4.2 Construction of the lexicon spectrum

The general scheme is the interpretation, as fuzzy sets of the lists obtained for every speaker/shared field of experience (i.e., a spectrum in the sense proposed here), and the subsequent aggregate of that information in order to obtain the lexical spectrum of the shared field of experience. This stepped constructive process enables not only the integration of further information but also the election of diverse mechanisms of integration, as shown below. Although we suggest here those which we consider more adequate, other options are open to adapt the model to different conditions.

A compatibility of the lexicon with the prototype supplied by the shared field of experience is obtained and can be interpreted as the availability index of said shared field of experience (Chart 1).

The values obtained for a token, and their gradation, are simple to interpret:

1. Close to one: the token belongs to the nucleus of the shared field of experience and is accessible for all speakers; i.e., it is very likely that any speaker produces it during the availability trial.

© 2014. John Benjamins Publishing Company
All rights reserved

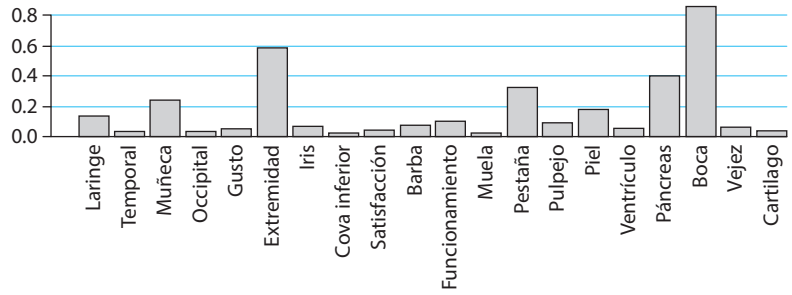


Chart 1. Structure of a shared field of experienceas a gradation of compatibility.
(Note: for English equivalents to horizontal axis items, see 4.2)

2. Close to zero: the token belongs to the periphery and is either accessible only by a segment of the speakers, or is a concept that has little in common with the prototype. It was impossible to discover from the results of the trials why a term diverges from a shared field of experience.

The process begins with the interpretation of the trial results by a certain speaker for a particular shared field of experience, through a fuzzy set that represents the linguistic performance accomplished by this individual in the lists of words produced.

Basing our study on the hypothesis that the order of lexical items on the lists supplied by each speaker is relevant with respect to the availability of each term – at least in an initial approximation – a representation of availability for each list of terms is built following a descending law with regard to its position. For this purpose we have chosen a Zipf-Mandelbrot law, which is widely used for linguistic studies, states that in a corpus of natural language, the frequency of any term is inversely proportional to its rank in the order of its frequency of occurrence.

Therefore each term for each speaker and shared field of experience is assigned a value (t) according to two parameters and the Zipf-Mandelbrot scheme: the position occupied on the list (n) and a constant for each problem (k):

$$t = \frac{k}{n} \tag{2}$$

Thus, for example, from the following list of terms:

> *boca* 'mouth', *extremidad* 'extremity', *páncreas* 'pancreas', *pestaña* 'eyelash', *muñeca* 'wrist', *piel* 'skin', *laringe* 'larynx', *funcionamiento* 'function', *pulpejo* 'fleshy part', *barba* 'beard', *iris* 'iris', *vejez* 'old age', *gusto* 'taste', *ventrículo* 'ventricle', *satisfación* 'satisfaction', *cartílago* 'cartilage', *occipital* 'occipital', *temporal* 'temporal', *muela* 'molar', *cava inferior* 'lower vena cava'

© 2014. John Benjamins Publishing Company
All rights reserved

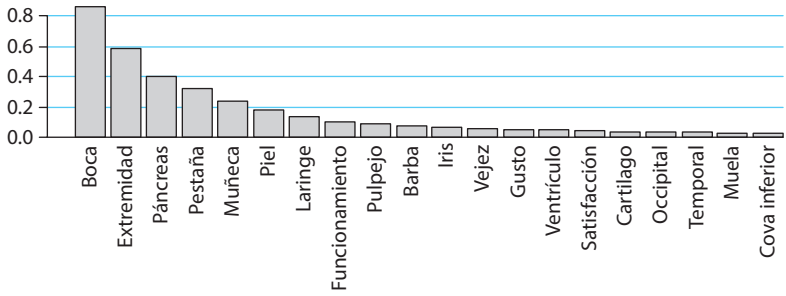


Chart 2. Representation of the lexical spectrum of Center of interest 1. The Human Body

we will obtain a graphic representation of the individual spectrum expressed by the individual during the trial as shown in the following Chart 2.

On the assumption that the structure of a shared field of experience is a common construct in the lexical structures of each speaker, the relevance of each term among the different speakers is incorporated by means of an additive law. Here it is about accumulating information, but not directly adding this up; i.e., the information already gathered is not decreased by new information. There are many operators that would satisfy the conditions needed, but in view of the lack of data, we chose probability addition as the simplest solution: if *a* and *b* are two compatibilities of a term for two trials, the resulting compatibility is calculated as follows:

$$a + b - a \cdot b \tag{3}$$

whose characteristic is that the occurrence of a term always adds more relevance to the term, in direct ratio with the relevance of its occurrence. In Chart 3 below a simulation shows what would happen when information provided by three speakers (three iterations) is incorporated.

These iterations were developed from the following values:

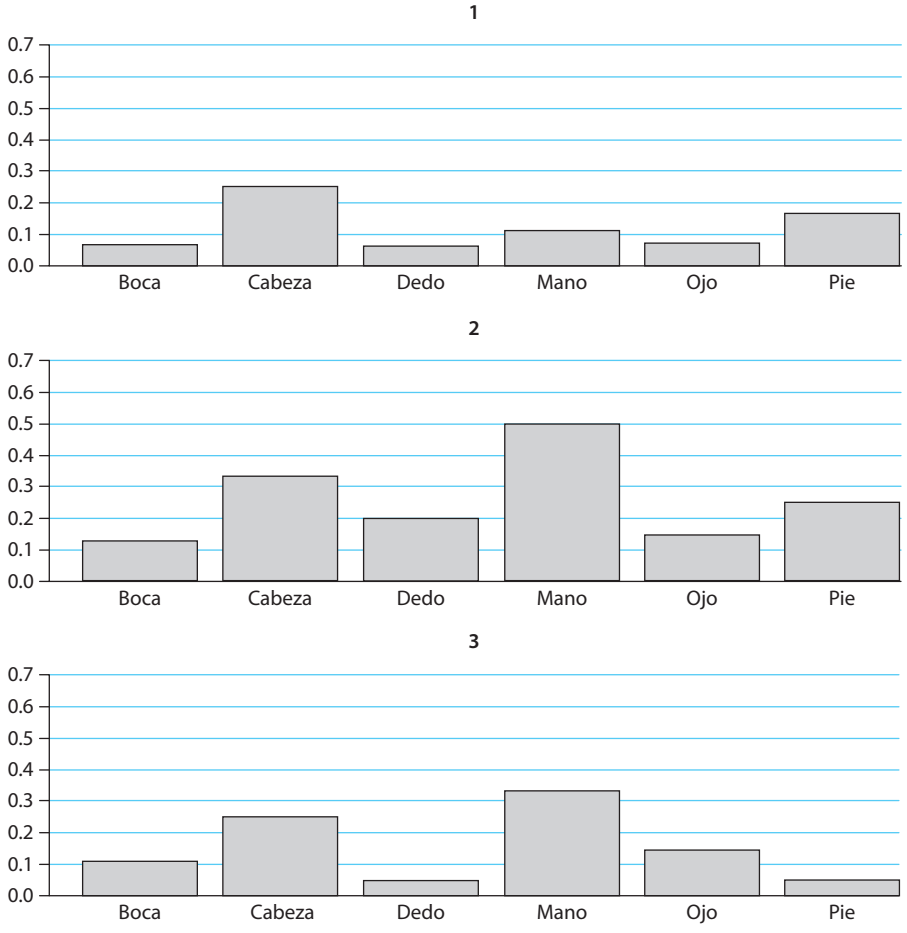
	boca **'mouth'**	***cabeza*** **'head'**	***dedo*** **'finger'**	***mano*** **'hand'**	***ojo*** **'eye'**	***pie*** **'foot'**
Subject 01	0.0666	0.25	0.0588	0.1111	0.0714	0.1666
Subject 02	0.125	0.3333	0.2	0.5	0.1428	0.25
Subject 03	0.1111	0.25	0.0476	0.3333	0.1428	0.05

The first step was to incorporate the information supplied by subjects 1 and 2 using the rule of probability discussed above and these were the results obtained:

© 2014. John Benjamins Publishing Company
All rights reserved

mouth: 0.0666 + 0.125 – 0.0666 · 0.125 = 0.1833
head: 0.25 + 0.3333 – 0.25 · 0.3333 = 0.4999
finger: 0.0588 + 0.2 – 0.0588 · 0.2 = 0.2470
hand: 0.1111 + 0.5 – 0.1111 · 0.5 = 0.5555
eye: 0.0714 + 0.1428 – 0.0714 · 0.1428 = 0.2040
foot: 0.1666 + 0.25 – 0.1666 · 0.25 = 0.3749

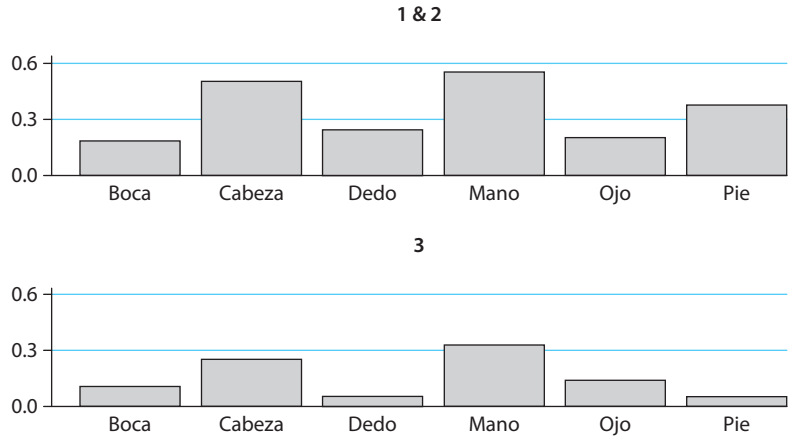
Iteration 1: evaluations of productions by each speaker (1, 2 and 3) are constructed[11]



11. The terms have been reorganized so that they are in the same position in all the figures. Order of occurrence in the lists is already taken into consideration in their compatibility assessment.

© 2014. John Benjamins Publishing Company
All rights reserved

Iteration 2: information supplied by speakers 1 and 2 is 'accumulated'



Iteration 3: information supplied by speaker 3 is incorporated into the information already gathered

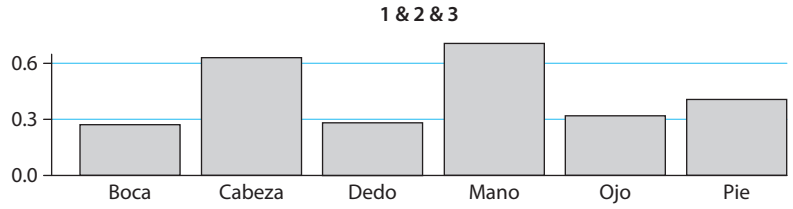


Chart 3. Incorporation process of individual spectra in shared field of experience 1. The Human Body

Thus the start values were structured as follows:

	boca **'mouth'**	***cabeza*** **'head'**	***dedo*** **'finger'**	***mano*** **'hand'**	***ojo*** **'eye'**	***pie*** **'foot'**
Subject 1&2	0.1833	0.4999	0.2470	0.5555	0.2040	0.3749
Subject 3	0.1111	0.25	0.0476	0.3333	0.1428	0.05

Incorporation of information supplied by subject 3 to the accumulated data gave the following results:

mouth: $0.1833 + 0.1111 - 0.1833 \cdot 0.1111 = 0.2740$
head: $0.4999 + 0.25 - 0.4999 \cdot 0.25 = 0.6249$
finger: $0.2470 + 0.0476 - 0.2470 \cdot 0.0476 = 0.2829$
hand: $0.5555 + 0.3333 - 0.5555 \cdot 0.3333 = 0.7037$

© 2014. John Benjamins Publishing Company
All rights reserved

eye: $0.2040 + 0.1428 - 0.2040 \cdot 0.1428 = 0.3177$
foot: $0.3749 + 0.05 - 0.3749 \cdot 0.05 = 0.4062$

and the resulting values were:

	boca **'mouth'**	***cabeza*** **'head'**	***dedo*** **'finger'**	***mano*** **'hand'**	***ojo*** **'eye'**	***pie*** **'foot'**
Subject 1,2&3	0.2740	0.6249	0.2829	0.7037	0.3177	0.4062

Specifically, in order to incorporate the information supplied by subjects 1 and 2 for the term *boca* 'mouth', the following operation was carried out:

$$0.0666 + 0.125 - 0.0666 \cdot 0.125 = 0.1833$$

and in turn to incorporate the information supplied by subject 3 to that already accumulated by the other two:

$$0.1833 + 0.1111 - 0.1833 \cdot 0.1111 = 0.2740$$

and so successively for all the terms and all the subjects.

Accumulating information produces a continuous increment effect in compatibility values between terms and the shared field of experience. This may lead to excessive evaluation of the compatibility with the concept expressed by the shared field of experience. Thus a control mechanism is needed in order to obtain the adequate evaluation of compatibility able to respond to the requirements of the intended representation. The proposed mechanism for our study is the determination of the parameter k in (2). This value enables us to scale the initial evaluation for every speaker in order to carry out the intended control.

But this mechanism brings about another problem: which is the adequate criterion to obtain a satisfactory representation? We have assumed that the best informative model is such that just one term shows perfect compatibility with the shared field of experience (i.e., obtaining evaluation indistinguishable from 1). Nevertheless, as pointed out above, another approximation may be perfectly correct and justifiable, for example the Maximum Entropy Principle in some of its different versions (De Luca & Termini, 1972). However, this last approximation would introduce a symmetric distribution of compatibility values, with numerous elements gathered both to 0 and 1, and the remainder in between, thus providing a maximum of segregation of the compatibility values.

Chart 4 shows an example of the result of this process in a real case where the terms have been ordered according to their relevance. (From left to right: *head, nail, tibia, vein, femur, nape, cheek, beard, biceps, bilis, aorta, carpus, obese, tarsus, nipple, dream, pleura, dorsal, pubis.*)

© 2014. John Benjamins Publishing Company
All rights reserved

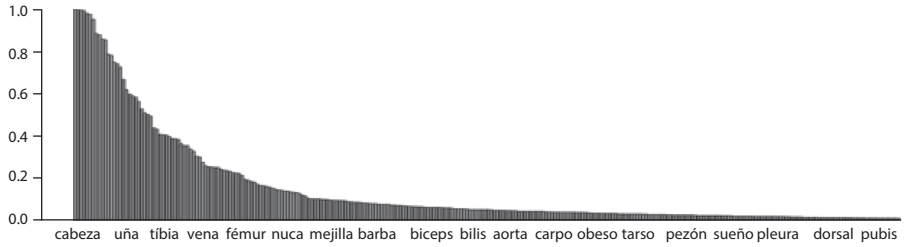


Chart 4. Ordered incorporation of individual spectra in Center of Interest 1. The Human Body

5. Conclusions

Representation of data extracted from word availability trials has revealed the early existence of some problems and inadequacies not solved at the time, not even by means of the complex theoretical models that some linguists have developed (Apresjan 1995, 2000; Mel'čuk, 1995; Miller & Fellbaum, 1991). Our contribution intends to build up a framework for representing that information and attaining a clearer understanding of the whole process involving word availability and the trials designed to test it.

A fresh start has been made by attempting to use the simplest functional hypothesis possible, and the most reliable framework that we have found to express the model obtained has been the Fuzzy Sets Theory.

In particular, the hypotheses, contrasted in the course of the research, are as follows:

- Extremely available words will be produced by many speakers and appear in initial positions of the lists.
- Poorly available terms, as they are not commonly used nor connected with the shared field of experience, will tend to be produced by few speakers and/or in final positions of the lists.
- Speakers from the same population tend to share a similar lexicon, since they are affected by the same socio-cultural factors. Therefore, by integrating the obtained information, availability for a particular shared field of experience may be recreated. The association token-availability for the produced terms will be referred to as *lexical spectrum*.

From the above hypotheses, the following process is proposed in order to determine the availability:

© 2014. John Benjamins Publishing Company
All rights reserved

- An initial evaluation of the availability is carried out for every trial, i.e., for every speaker and shared field of experience. This initial evaluation would incorporate any accessible information as, for instance, the timing during the trial. Since in our project we only had access to the word position in the lists, we chose a Zipf-Mandelbrot law based in the position of every term in the list because of its slow, deliberate descending behavior.
- Initial evaluations are accumulated following an additive law; that is to say, every time a word is found, the evaluation gathered rises in direct proportion to the incorporated availability. A probabilistic sum law has been chosen for this purpose.
- A control mechanism is necessary to cope with the increase of the availability values in the accumulating process. This mechanism should respond to specific criteria for representing information. The proposed mechanism has been the use of the factor *k* in the Zipf-Mandelbrot law for the initial evaluation of the trials. The criterion that we have chosen specifies that just one term must be fully available in the shared field of experience; that is to say, with evaluation indistinguishable from unity. This criterion leads to obtaining a suitable value of *k*.

In addition to the correspondence with the proposed conceptual framework, representing the linguistic spectra as fuzzy sets provides an extensive and useful set of tools (Zimmermann, 2001). For example:

1. Identifies the differences in the vocabulary between different shared fields of experience or in the same shared field of experience between different population groups, using difference and fuzzy symmetric difference operators;
2. Determines the size of the lexicon of a shared field of experience: fuzzy cardinality;
3. Determines the relevance value from which a lexical element may be considered to be in the nucleus of a shared field of experience: Fuzzy Expected Value (FEV).

In the seventies the Fuzzy Set Theory was widely accepted for treating linguistic problems. In this work we propose to recover this framework which we have found to be useful and flexible to redesign problems such as that of determination of word availability.

Lastly, but perhaps not less important, this model makes it possible to develop constructions unconsidered until now, such as calculation of individual lexical capacity (Villena-Ponsoda, Ávila-Muñoz, & Sánchez-Sáez, 2012, pp. 409–432), as well as the construction of collective lexical spectra for different cognitive prototypes.

© 2014. John Benjamins Publishing Company
All rights reserved

References

Apresjan, J. (1995). Theoretical linguistics: Formal models of language and systemic lexicography. In I. H. Lee (Ed.), *Linguistics in the Morning Calm 3*: Selected Papers from SICOL-1992 (pp. 3–30). Seoul.

Apresjan, J. (2000). *Systematic lexicography*. Oxford: Oxford University Press.

Ávila-Muñoz, A. M., & Sánchez-Sáez, J. M. (2010). La disponibilidad léxica. Antecedentes y fundamentos. In Ávila-Muñoz, A. M., & J. A. Villena-Ponsoda (Eds.), *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis* (pp. 37–81). Málaga: Sarriá.

Ávila-Muñoz, A. M., & Sánchez-Sáez, J. M. (2011). La posición de los vocablos en el cálculo del índice de disponibilidad léxica. Procesos de reentrada en las listas del léxico disponible de la ciudad de Málaga. *ELUA*, 25, 45–74.

Benítez, P. (1992). Disponibilidad léxica en la zona metropolitana de Madrid. *Boletín de la Academia Puertorriqueña de la Lengua Española* (BAPLE), 1(1), 71–102.

Butrón, G. (1987). *El léxico disponible: Índices de disponibilidad*. Unpublished doctoral thesis. Río Piedras: University of Puerto Rico.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsyhcology*, 14(1), 177–208.

Carnap, R. (1952). Meaning postulates. *Philosophical Studies*, 3, 45–78.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. DOI: 10.1037/0033-295X.82.6.407

Collins, A. M., & Quilian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour*, 8, 240–247. DOI: 10.1016/S0022-5371(69)80069-1

Cuetos, F. (1991). *Psicología de la escritura*. Madrid: Editorial Escuela Española.

De Luca, A., & Termini, S. A. (1972). Definition of a nonprobabilistic entropy in the setting of Fuzzy Sets Theory. *Information and Control*, 20, 301–312. DOI: 10.1016/S0019-9958(72)90199-4

Dimitrijêvic, N. (1969). *Lexical availability*. Heidelberg: Julius Gross Verlag.

Ellis, A. W. (1982). Spelling and writing (and reading and speaking). *Normality and Pathology in Cognitive Functions*, 113–146.

Friedman, M., Schneider, M., & Kandel, A. (1989). The use of Weighted Fuzzy Expected Value (WFEV) in Fuzzy Expert Systems. *Fuzzy Sets and Systems*, 31, 37–45. DOI: 10.1016/0165-0114(89)90065-1

Galloso-Camacho, M. V. (2003). *El léxico de los estudiantes preuniversitarios en el distrito universitario de Salamanca (Ávila, Salamanca y Zamora)*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua.

Gómez-Devís, M. B. (2003). *La disponibilidad léxica de los estudiantes preuniversitarios valencianos: Reflexión metodológica, análisis sociolingüístico y aplicaciones*. Valencia: Universitá.

Gougenheim, G., Michéa, R., Rivenc, P., & Sauvageot, A. (1956). *L'élaboration du français élémentaire. Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris: Didier.

Haber-Schaim, U., Cross, J. B., Abegg, J., Dodge, J. H., & Walter, J. A. (1979). *Curso de introducción a las ciencias físicas*. Barcelona: Reverté.

© 2014. John Benjamins Publishing Company
All rights reserved

Hernández-Muñoz, N. (2006). *Hacia una teoría cognitiva integrada de la disponibilidad léxica: El léxico disponible de los estudiantes castellano-manchegos*. Salamanca: Universidad.

Holyoak, K. J. (2002). Psicología. In R. A. Wilson & F. C. Keil (Eds.), *Enciclopedia MIT de las Ciencias Cognitivas* (pp. 29–40). Madrid: Editorial Síntesis.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 39, 170–210. DOI: 10.2307/411200

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago & London: The University of Chicago Press. DOI: 10.7208/chicago/9780226471013.001.0001

Levelt, W. J. L. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223–232. DOI: 10.1016/S1364-6613(99)01319-4

Levelt, W. J. L. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, 98(23), 13464–13471. DOI: 10.1073/pnas.231459498

Levelt, W. J. L., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1–75.

López-Chávez, J., & Strassburger-Frías, C. (1987). Otro cálculo del índice de disponibilidad léxica. In Presente, y perspectiva de la investigación computacional en México. *Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada*. México: UNAM.

López-Chávez, J., & Strassburger-Frías, C. (1991). Un modelo para el cálculo de disponibilidad léxica individual. In La *enseñanza del español como lengua materna* (pp. 91–112). Río Piedras: Universidad de Puerto Rico.

López-Morales, H. (1973). *Disponibilidad léxica en escolares de San Juan. San Juan de Puerto* Rico: ms.

López-Morales, H. (1978). Frecuencia léxica, disponibilidad y programación curricular. *Boletín de la Academia Puertorriqueña de la Lengua Española,* 6(1), 73–86.

López-Morales, H. (1979). Disponibilidad léxica y estratificación socioeconómica. In *Dialectología y sociolingüística. Temas puertorriqueños* (pp. 173–181). Madrid: Hispanova de Ediciones.

López-Morales, H. (1999). *Léxico disponible de Puerto Rico*. Madrid: Arco/Libros.

Lorán, R. (1983). *Un índice de disponibilidad léxica*. Mayagüez: University of Puerto Rico.

Lorán, R., & López-Morales, H. (1983). Nouveau calcul de l'indice de disponibilité. Unpublished Original. Puerto Rico: University of Puerto Rico.

Mackey, W. F. (1971). *Le vocabulaire disponible du français*. Paris-Brussels-Montreal: Didier.

Mel'čuk, I. (1995). The future of the lexicon in linguistic description and the explanatory combinatorial dictionary. In I. H. Lee (Ed.), *Linguistics in the Morning Calm 3*: Selected Papers from SICOL-1992 (pp. 181–270). Seoul.

Miller G., & Fellbaum C. (1991). Semantic networks of English. *Cognition*, 41, 197–229. DOI: 10.1016/0010-0277(91)90036-4

Romero-Rubilar, C. (2000). Variabilidad y prototipicidad en vocabulario disponible. *Revista de Lingüística Teórica y Aplicada*, 38, 139–148.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale: Lawrence Erlbaum.

© 2014. John Benjamins Publishing Company
All rights reserved

Sartori, G., Coltheart, M., Miozzo, M., & Job, R. (1994). Category specificity and informational specificity in neuropsychological impairment of semantic memory. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 537–554). Cambridge/London: MIT Press.

Troyer, A. K., Moscovitch, M., & Wincour, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology*, 11, 138–146. DOI: 10.1037/0894-4105.11.1.138

Urrutia, M. A. (2001). *Edad y acceso al léxico*. Unpublished master's thesis. Universtity of Concepción.

Villena-Ponsoda, J. A., Ávila-Muñoz, A. M., & Sánchez-Sáez, J. M. (2012). Patrones sociolingüísticos del vocabulario disponible. Condicionamiento estratificacional de la capacidad léxica en la ciudad de Málaga (Proyecto CONSOLEX). In *Tradición y progreso en la lingüística general* (pp. 409–432). Valladolid: Universidad.

Wittgenstein, L. (1953). *Philosophical investigations*. New York: McMillan.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.

Zimmermann, H. J. (2001). *Fuzzy Set Theory and its applications*. Boston: Kluwer Academic Publishers. DOI: 10.1007/978-94-010-0646-0

Web references

<http://www.linguas.net/Proyectos/LexiDisp/tabid/73/Default.aspx> [last consulted: 16 March 2013]

<http://www.dispolex.com/> [last consulted: 16 March 2013]

<http://www.linguas.net/portalpreseea/Metodolog%C3%ADa/tabid/474/Default.aspx> [last consulted: 26 March 2013]

Authors' addresses

Antonio M. Ávila-Muñoz and José María Sánchez-Sáez
Department of General Linguistics
Humanities Faculty
Teatinos Campus
Universidad de Málaga
29071 Málaga (Spain)

amavila@uma.es jmss70@gmail.com

About the authors

Antonio Manuel Ávila-Muñoz, Ph.D. in Spanish Studies and Linguistics, is a Senior Teacher and Lecturer of General Linguistics at Malaga University (UMA), Spain. As a researcher, he has taken part in several national and international research projects. His interest in Sociolinguistic studies has mainly been focused on works and publications about several sub-disciplines such as social dialectology, linguistic variation and cognitive sociolinguistics. He has published several books about sociolinguistics and lexical analysis, apart from specific works in periodical publications and chapters in different books.

© 2014. John Benjamins Publishing Company
All rights reserved

José María Sánchez-Sáez, Ph.D. in Mathematics (2011), University of Malaga. Contributor to research group since 1991. His main areas of academic interests are Statistics and Operational Research, both theoretical and applied developments. He also works in other fields, like Fuzzy Sets Theory, as can be seen in this work.

© 2014. John Benjamins Publishing Company
All rights reserved