

Lexical Selection, or How to Bridge the Major Rift in Language Processing

WILLEM J. M. LEVELT, Nijmegen

1. Introduction

The claim that a language's mapping of sound to meaning is (by and large) arbitrary needs no defense nowadays. But it would be false to conclude that the claim is a trivial one. It isn't. An explicit theory of language should give an account of this arbitrariness. Where in the system is it to be located? Is it limited to a single module, such as the lexicon, or is it distributed all over the system? Are there boundary conditions on arbitrariness? If so, are they universal?

And additional questions should be raised by those (like me) who consider language to be a mental faculty: What kind of learning mechanism is involved in the acquisition of arbitrary sound/meaning relations? Is this arbitrariness reflected in language processing, in particular in the processes of speaking and of language comprehension? Is it, in particular, the case that processing is different where relations are arbitrary than where they are systematic or rule-governed?

It is this latter question that will be the focus of the present paper. After some introductory remarks about what I will call the major rift in the system, the main locus of arbitrariness, I will discuss two central properties of lexical processing, activation and selection. I will then argue that the relation between activation and selection is a different one across this major rift than it is at either side of it. The claim will be supported by data on lexical access in speech production.

2. Interfacing: systematicity and arbitrariness

Mapping meaning to sound involves at least the following three interfaces: meaning to syntax, syntax to phonology, and phonology to phonetics. There is arbitrariness in each of these interfaces, but to different degrees. Although theories vary substantially in the way they represent the meaning-to-syntax mapping, all recognize systematicity in the way semantic arguments are mapped onto syntactic functions. Semantic ar-

guments can be universally ordered on a scale of saliency, ranging from human agents via themes to sources, goals and a host of minor thematic roles (FILLMORE 1977). Grammatical functions, in their turn, can be ordered on a scale of prominence (KEENAN 1976), ranging from subject, via different types of object to obliques. And the systematicity is in the most salient argument going for the most prominent function. If there is a human agent, it will preferably be mapped on the subject function. If that slot happens to be occupied, it will go for the next function in the prominence hierarchy, etc.

This preference also shows up in the lexicon. There is a canonical order in which lexical items map semantic arguments onto grammatical functions, and it follows the same systematicity. Most verbs that have an agent as semantic argument (such as *give*) will map it onto their external (subject) function, etc. But this canonical order is often violated in the lexicon (such as in *receive*). Moreover, most verbs allow for two or more different mappings (such as actives and passives). Also the sheer number of grammatical functions that a verb requires may differ from the number of semantic arguments it expresses (such as in raising verbs).

Since there are many more different semantic arguments than there are syntactic functions, syntax cannot fully absorb the wealth of semantic distinctions. And the resulting mapping is often quite arbitrary. Neither can syntax absorb the richness of semantic modification. Syntax is rather more like a Procrustean bed that forces unequals to become equal. Syntax, one could say, is the poor man's semantics.

Turning now to the sound-related side of the system, the phonology-to-phonetics mapping, we find a mirror image of the latter situation. There is substantial systematicity in this mapping as well, but now it is phonology that cannot absorb the richness of phonetics. Articulatory gestures and their acoustic effects can range continuously where phonological representations are discrete. The same phonological distinction can usually be realized in an unlimited number of ways. Not only do speakers of the same language differ in their articulatory realizations of a phonological pattern in rather arbitrary ways, but the same speaker varies considerably in the way phonological features are physically realized, dependent on phonetic context, key, register, rate and formality of speech. This variability is not always systematic; it can, in fact, be quite arbitrary as well.

There is, it should be added, also an inverse indeterminacy. Underlying phonological distinctions may get lost in the phonetic signal; there is phonological reduction all over the place in normal fluent speech. It is, therefore, probably correct to characterize the phonology-to-phonetics relation as a some-to-many mapping. Or, phonology is the poor man's phonetics. The poor man, of course, is the language user who has to

remember the sound forms of his language. It is impossible to store the infinite range of well-formed articulatory patterns. But it *is* possible to store a finite number of articulatory *tasks* (such as to close the lips, or to raise the velum) that the articulatory system will have to execute in order to realize the language's sound distinctions. The execution of each such task is a one-to-many mapping that is not stored, but the natural product of an intelligent motor system (BROWMAN and GOLDSTEIN 1990) that varies rather arbitrarily between and within speakers.

The most significant arbitrariness in the system, however, resides in the syntax-to-phonology mapping. Still, this arbitrariness is restricted in locus. There is, for instance, great systematicity in the way syntactic constituent structure is reflected in phonological constituent structure. In many languages, for instance, the lexical head of a syntactic phrase becomes the completion of the current phonological phrase. Intonational phrase boundaries tend to coincide with clause boundaries, etc.

The major rift in the system is internal to the lexicon. It is in the way morphemes, as meaningful syntactic units, map onto phonological patterns. Apart from homophony (which can be substantial, like in Chinese), the mapping is by and large one-to-one but almost completely arbitrary. There is no reason why a cat should be called /kæt/; it is just an accident of English. There is no rule or systematicity by which this fact can be predicted. Only large-scale statistical analyses of the lexicon show that there is some systematicity even here. KELLY (1992) reviews some of it. In English, for instance, nouns tend to contain more syllables than verbs, and different from verbs nouns tend to have word accent on the first syllable. Nouns also contain front vowels more often than verbs do (SERENO and JONGMAN 1990), etc. KELLY argues that the listener may be using such statistical regularities in parsing. But there is just no way for a listener to access the lexicon on the basis of these regularities alone; they are really quite marginal to the system.

In short, we are observing an almost Cartesian state of affairs. There is, on the one hand, a meaning/syntax system with fairly systematic internal relations. There is, on the other hand, a phonology/phonetics system with fairly systematic internal relations. But as far as these two are connected via the lexicon, that connection is as arbitrary as the pinal gland. It is, then, a reasonable question to ask, whether linguistic processing reflects this situation. In particular, is this major rift apparent in lexical processing?

3. Lexical activation and selection in production

3.1 The lexical network

The notion of activation spreading has always been around in theories of lexical access. Since COLLINS and QUILLIAN (1969) activation spreading through lexical networks has become a major theoretical device in the study of lexical access. SCHNELLE (1989, p.167) is right in stating that "phonetic, phonological, and morphological data connected with the problem of *lexical access* (...) and the *phenomena of speech production* (...) provide the best problem areas with which to start" (i.e., to start the modelling of parallel linguistic processing). Here I will take up this challenge, and give a short outline of our model of lexical access in speech production. I will then discuss some data on lexical activation and selection that are relevant to the rift issue introduced above.

The model was largely developed by ROELOFS (1992) as a solution of the so-called "hyperonym problem" formulated in LEVELT (1989). A further introduction to the model can be found in BOCK and LEVELT (in press). In the model the production lexicon is represented as a network through which activation can spread. It is not a connectionist network, but one in the tradition of COLLINS and QUILLIAN (op. cit.). That is, both nodes and arcs are labelled entities, and there may be various conditions on the spreading of activation between nodes.

Figure 1 (see next page) represents a tiny part of the production lexicon. A lexical item is represented by a triple of connected nodes. Each node resides at a different stratum. The top stratum is the conceptual level. Nodes represent concepts, and arcs the relations that hold among them. The notion of a cat is represented by the node CAT, and its meaning is represented by the network of relations to other conceptual nodes. There is, for instance an *is-a*-relation to the node ANIMAL, a *food*-relation to MEAT, etc. A conceptual node can be activated by activation spreading through the network. For instance, if CAT is an active node, some of its activation will spread to the node DOG via connecting arcs. Also perceptual information, such as seeing a cat, may activate the corresponding node.

Some conceptual nodes have a direct arc connection down to the next level, which is called the lemma level. CAT, for instance, is directly connected to the lemma node "cat". Such concepts are "lexical concepts", i.e., concepts for which there is an entry in the lexicon. The lemma level is a syntactic stratum. The arc connections represent a lemma's syntactic properties. The lemma *cat* is of syntactic category noun. German *Katze* is, in addition, of female gender. Similarly, subcategorizations of verbs can be represented as network relations at this level, etc. One important

source of lemma activation is an active lexical concept. But in addition lemmas can be activated by the spoken or printed word that corresponds to the lemma. Finally, a small set of lemma nodes (not represented in Figure 1) can be activated by stratum-internal syntactic activation. Among them are a language's closed class items.

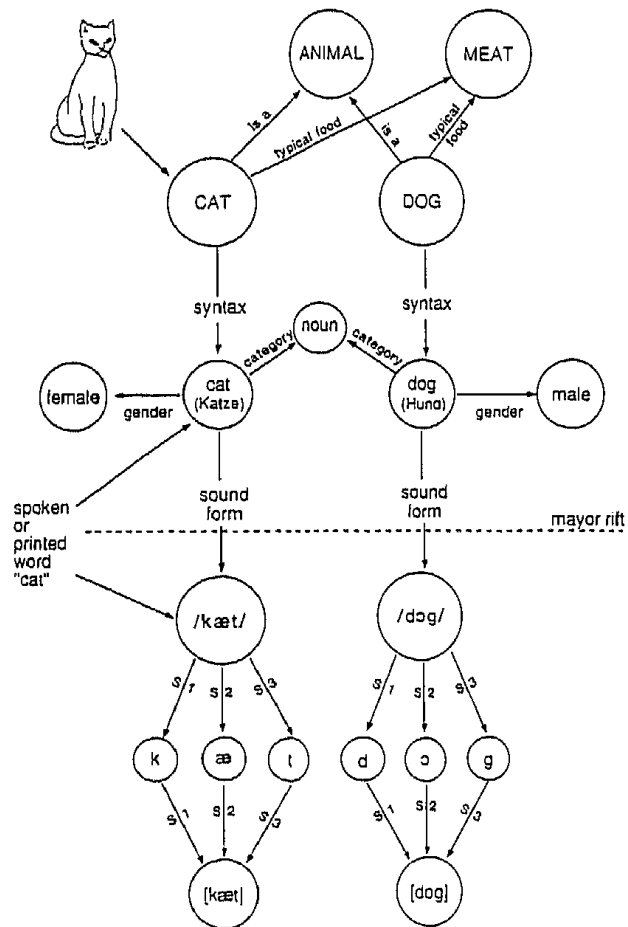


Figure 1: Fragment of a lexical production network

Each lemma node has an arc connection to a node at the bottom stratum, a so-called lexeme node. Lexeme nodes represent a lexical item's

form properties by way of a network of labelled relations to various form nodes. There are, on the one hand, nodes that represent an item's phonological segments (probably at a rather abstract or "underspecified" level). And there are, on the other hand, nodes that represent a word's metrical or foot structure (not shown in the figure). Finally, there may be a level of syllable nodes here. In LEVELT (1992) I have argued that speakers may have a mental syllabary, a store of phonetic descriptions for the not too infrequent phonological syllables in their language. These phonetic descriptions are in terms of the articulatory tasks mentioned above. Lexeme nodes can be activated by their corresponding lemma nodes, but also when the corresponding word is heard or read.

3.2 Lexical selection

Lexical selection in production is the choice of a lemma. Selectional errors such as "... *carrying a bag of cherries. I mean grapes*" (STEMBERGER 1985) can be explained from activation spreading at the conceptual level. When CHERRY is an active node, then the closely related GRAPE will become coactivated, and it will spread some of its activation down to the lemma level. ROELOFS (1992) modelled lexical selection in the following way: The probability of choosing a particular lemma (call it the target lemma) at any one discrete time interval is the ratio of the target's activation to the total activation of all lemmas in the response set (this is the so-called Luce ratio). So, any activated lemma in that set has a non-zero probability of being selected. That, apparently, happened to *cherries* (or rather *cherry*) in the above error.

But ROELOFS's primary empirical evidence is not error data but reaction time data, in particular word onset latencies in picture naming. His main experimental task is an interference paradigm. The subject has to name a picture, but at some moment a visually presented distracter word appears, which the subject has to ignore. Usually, such distracters affect the picture naming latency. And that is especially so for semantically related distracters. If, for instance, a picture of a cat is presented and simultaneously the word "dog" is flashed, the naming response "cat" is delayed. This, at least, happens when the subject knows that there could as well be a picture of a dog (i.e., "dog" is in the response set). The reason for the delay is that the Luce ratio for the target ("cat") will be smaller when the semantic alternative ("dog") gets extra activation. The model gives a precise quantitative account of both the interference data in the literature and of newly acquired data.

3.3 Activation spreading and the rift

These and similar experiments (LEVELT et al. 1991) show that activation can spread freely from the conceptual to the lemma level. The target concept spreads its activation to related concepts. These, in turn, spread their activation to their lemmas. Lexical selection is the outcome of a competition between coactivated lemmas.

One would now expect that the same story should *mutatis mutandis* hold for the relation between the lemma level and the form level: Any active lemma will spread some of its activation to its lexeme node, and the most highly activated lexeme has the best probability of being selected. And that is precisely what connectionist models (such as DELL 1986 or MACKAY 1987) predict. But then one doesn't reckon with the rift.

In order to test whether any activated lemma spreads its activation to the lexeme level, LEVELT et al. (1991) devised the following experimental procedure. The task is again a naming task. The subject names one picture after another. In about one third of the trials an additional stimulus is presented to the subject, but it is not a distracter stimulus. The stimulus can be a spoken word (like "house") or a non-word (like "sef"). The subject's (secondary) task is to decide whether the acoustic probe is a word or a non-word. The decision is indicated by pushing a "yes" button or a "no" button as fast as possible. And, of course, the subject has to name the picture. In the critical experiment, the acoustic probe began (on average) at 73 milliseconds after the picture appeared. From earlier experiments we knew that this was the right moment to measure activation of the lexeme, i.e., phonological activation of the target word.

How could this phonological activation be measured? This was done by presenting as acoustic probe a word that is phonologically related to the target, i.e., to the name of the picture. For instance, when the picture was one of a cat, the acoustic probe could be "cap", and the subject would push the "yes" button because "cap" is a word. It turns out that the lexical decision to "cap" is slower than the lexical decision to a phonologically unrelated word (such as "pill"). In our experiment this difference (between phonologically related and unrelated probes) amounted to a highly significant 88 msec. It indicates that the lexeme node of the target word is highly active at the moment of measurement. Or in other words, the target lemma spreads its activation across the rift.

But what about alternative lemmas? Will semantically related lexical items become phonologically active as well? For instance, if the subject is naming the cat's picture, will there be phonological activation of "dog"? We know from ROELOFS' work that lemmas that are semantically related to the target will become active as well. This was moreover confirmed in

our lexical decision paradigm. When we presented "dog" as a probe word when the target was "cat", lexical decision was substantially slowed down (as compared to a neutral probe like "pill"), on average by 106 msec. Will such a coactivated lemma spread its activation over the rift? This could be tested by using a lexical decision probe that is phonologically related to this semantic alternative. If the alternative is "dog", the probe could be "dot". Would lexical decision to such probes be slowed down (as compared to neutral probes)? Our experiment showed that this was not the case. Actually, the lexical decision latencies were on average 2 msec. faster. There was not the slightest indication that coactivated lemmas spread any of their activation to their lexeme nodes.

4. Conclusion

This surprising finding leaves us with the following conclusion: Only *selected* lemmas can spread their activation over the rift, merely *activated* lemmas don't. In speech production lexical selection is apparently a necessary condition for initiating the encoding of sound form. It also means that the mechanism of phonological encoding is not like the mechanism of lexical selection. It is not the case that there is a competition between alternative active lexemes, one of which becomes selected (following Luce's rule). Rather, only a single lexeme, the one corresponding to a selected lemma, becomes activated.

What we have learned, following SCHNELLE's challenge, is that there is no unlimited cascading of activation through the lexical network. There is a rift in the middle, and processing at the two sides follows different principles. Isn't this what FODOR (1983) called modularity?

References

- Bock, J. K. and Levelt, W. J. M. (in press): "Language Production: Grammatical encoding". In: M. Gernsbacher (Ed.): *Handbook of psycholinguistics*. Academic Press, New York.
- Browman, C. P. and Goldstein, L. (1990): *Gestural specification using dynamically-defined articulatory structures*. Haskins Laboratories Status Report on Speech Research, SR-103/104, 95-110.
- Collins, A.M. and M. R. Quillian (1969): "Retrieval time from semantic memory". *Psychological Review* 82: 407-428.
- Dell, G. S. (1986): "An activation spreading of theory of retrieval in sentence production." *Psychological Review* 93: 283-321.
- Fillmore, C. J. (1977): "The case for case reopened." In: P. Cole and J.M. Sadock (Eds.): *Syntax and semantics: Vol. 8. Grammatical relations*. Academic Press, New York.
- Fodor, J. (1983). *The modularity of mind*. M.I.T. Press, Cambridge, MA.

- Keenan, E. L. (1976). "Toward a universal definition of 'subject'." In: C. N. Li (Ed.): *Subject and topic*. Academic Press, New York.
- Kelly, M. H. (1992): "Using sound to solve syntactic problems: The role of phonology in grammatical category assignments." *Psychological Review* 33: 349-364.
- Levelt, W. J. M. (1989): *Speaking: From intention to articulation*. M.I.T. Press, Cambridge, MA.
- Levelt, W. J. M. (1992): "Accessing words in speech production: stages, processes and representations." *Cognition* 42: 1-22.
- Levelt, W. J. M., H. Schriefers, D. Vorberg, A. S. Meyer, T. Pechmann, and J. Havinga (1991): "The time course of lexical access in speech production: A study of naming." *Psychological Review* 98: 122-142.
- MacKay, D. (1987): *The organization of perception and action: A theory for language and other cognitive skills*. Springer-Verlag, New York.
- Roelofs, A. (1992): "A spreading activation theory of lemma retrieval in speaking." *Cognition* 42: 107-142.
- Schnelle, H. (1989): "The challenge of concrete linguistic description: Connectionism, massively parallel distributed processing, net-linguistics." In: H. Schnelle and N. O. Bernsen (Eds.): *Logic and linguistics. Research directions in cognitive science. European perspectives. Vol.2*. Lawrence Erlbaum, Hove.
- Sereno, J. A. and A. Jongman (1990): "Phonological and form class relations in the lexicon." *Journal of Psycholinguistic Research* 19: 387-404.
- Stemberger, J. P. (1985): "An interactive action model of language production." In: A. W. Ellis (Ed.): *Progress in the psychology of language (Vol.1)*. Lawrence Erlbaum, Hillsdale, N.J.