Metrical Segmentation and Lexical Competition: A Happy Affair?

Beatrice de Gelder (1,2) and Jean Vroomen (1)

Tilburg University (1)
Université Libre de Bruxelles (2)

There is nothing language specific about what it is listeners are doing, i.e. understanding words. So, one might argue, there should not be something language specific about the way they do it. But, universality and language specificity are two sides of the same coin. Understanding specificity requires searching for generality, just like looking for universals consists of studying language-specific manifestations. Linguistic concepts like those of the syllable, feature, phoneme, or mora are tools for understanding linguistic reality in this double aspect of specificity and universality. Like in any science, one should not simply judge a good theory in terms of degree of match between its concepts and behavioral facts, as if the former were merely names for the latter that label underlying realities. We insist briefly on this point because the benefit of interlinguistic comparisons would be impoverished if research would simply conclude that in one language the processing units are one type of segment, in another language another type of segment, etcetera. This would in the end not give us a real understanding of linguistic diversity and universality. It would only list what is on the colour chart and not what is on the painting. In this spirit, the motive of the following comments must not be understood as one of figuring out which of two (or more) alternative accounts is the right one and which the wrong one. On the other

hand, too much agreement, like too much of its opposite, means that one might be asking wrong questions.

Competitions effects: what are they?

Cutler, Norris and McQueen (1994) have made a strong case for a prelexically operating strategy in which native listeners of English have their lexical access attempts guided by the notion that strong syllables are good cues to word beginnings. Since Dutch has most of the properties that make the MSS a good strategy for English, we expected, and also found, a very similar picture for native listeners of Dutch, as well as for those who acquired Dutch as second language (van Zon & de Gelder, 1993; Vroomen, van Zon & de Gelder, in preparation; de Gelder, van Zon & Vroomen, in preparation). Most recently, Cutler et al. (1994) propose a combination of two types of processes: a rhythm-based strategy whereby the speech stream is segmented in relevant units for the purpose of word recognition and a properly lexical strategy where different words compete for recognition. Likewise, in recent studies in our own laboratory, we obtained evidence for lexical competition. We conducted two cross-modal priming experiments in which the effects of competition on spoken word recognition were investigated (Vroomen & de Gelder, submitted). The subjects heard Dutch CVCC (e.g., *melk*, milk) or CVC (e.g., *bel*, bell) words embedded in bisyllabic nonsense words. The second syllable of this nonsense word was either weak (*melkem* and *belkem*) or strong, and the cohort size of competitors starting with strong syllables was either small (*melkeum* and *belkeum*) or large (*melkaam* and *belkaam*). The cohort-size of the weak syllables is effectively zero, as there are no words in Dutch that start with an unvoiced stop consonant followed by schwa. These auditory nonsense words served as primes for a visual target (MELK or BEL). At an inter-stimulus interval (ISI) of 250 msec, *melkem* had the largest facilitatory effect on the target MELK, *melkeum* had an intermediate effect, and the facilitatory effect of *melkaam* was smallest. There was no difference in the facilitatory effects of *belkem*, *belkeum* and *belkaam* on BEL. These results are strikingly similar to those reported by Cutler, Norris and McQueen (this volume) in that they show that the activation of a target word is proportionate to its number of competitors. Words with many

competitors (*melkaam*) are less activated than words with few competitors (*melkeum*), and these are, on their turn, less activated than words with no competitors (*melkem*). If there is no phonetic overlap, as with BEL in *belkem*, *belkeum*, and *belkaam*, there is also no competition effect.

In addition, the cross-modal priming paradigm allows one to trace the activation of target words in time as one can vary the ISI between prime and target. This is an important feature, because, if properly conceived, competition effects should have a time course. That is, before competitors can start to inhibit a target word, they must first be sufficiently activated. The prediction is thus that competition effects will disappear if the ISI is short. And indeed, that is exactly what was found. At an ISI of 0 msec, large facilitatory effects for all primes were observed, but there was no difference any more between words with many, few, or no competitors. We interpret this finding as rather direct evidence for competition at the word level, and so far we thus agree with Cutler et al (1994).

But, the notion of competition is, as Cutler et al. (1994) note, ambiguous. There are at least two interpretations of competition. The first we would like to call *direct* competition. It is similar to the way competition is instantiated in TRACE (McClelland & Elman, 1986) and also SHORTLIST: words which are activated above a threshold decrease the activation of other words with which they overlap in time such that the "rich get richer" effect emerges. The other notion of competition is *indirect* and it is similar to competition in the Neighbourhood Activation Model (NAM) of Goldinger, Luce and Pisoni (1990). In NAM competition does not arise directly at the activation level of the word, but in a subsequent decision stage. Words in NAM are recognized if the ratio of the frequency-weighted input of a target word to its competitors is sufficiently high. If many competitors are activated, the ratio is lowered because the denominator is larger. In formula, this is:

$$p(\ identification_{target}\ ) = \frac{input_{target}\ x\ freq_{target}}{input_{target}\ x\ freq_{target} + \sum_{j=1}^{n} input_{competitor_j}\ x\ freq_{competitor_j}}\ 1$$

The data of Cutler et al. do *not* distinguish between direct and indirect

competition, because in word-spotting, reaction time is measured to a word which is recognized. Thus, both direct and indirect competition could have their effects on reaction time. Cross-modal priming is somewhat less ambiguous to interpret, because it is generally agreed that facilitatory effects reflect *pre*-activation of the target. So the idea is that cross-modal priming taps pre-decisional stages.

More direct evidence for the distinction between direct versus indirect competition comes from another study which we recently conducted. Direct and indirect competition have distinct predictions associated with each of them. Direct competition predicts that competition effects should be smaller in high-frequency (HF) targets, because HF targets are more able to inhibit their competitors such that the influence of their competitors is decreased (see Bard, 1991 for simulations with TRACE on this issue). Indirect competition predicts just the opposite: the chance of identifying a HF target decreases more than that of a LF target if there are more competitors. Take as an example a HF target with a frequency-weighted input of .6 whose competitors have a summed frequency-weighted input of either .5 (few competitors) or 1.0 (many competitors). The identification ratio's are then .6/(.6 + .5) = .54 with few competitors and (.6/.6 + 1.0) = .375 with many competitors. The competition effect for the HF target is thus .54 - .375 = .165. Now take a LF target with the same competitors, but whose frequency is twice as low as that of the HF target: the frequency weighted input is .3, and the identification ratio's are .3/(.3 + .5) = .375 for few competitors and .3/(.3 + 1.0) = .23 with many competitors. The competition effect for the LF target is now only .375 - .23 = .145, which is smaller than that of the HF target. Thus according to the *indirect* notion, competition effects should be smaller in LF targets, whereas these effects should be larger in the *direct* notion. We have recently tested these predictions by comparing competition effects in HF and LF targets. The results were exactly as predicted by *direct* competition: competition effects were largest in LF targets, and almost absent in HF targets. This pattern is stable, as we replicated it in a second study with new items and new subjects (Vroomen & de Gelder, in prep.). Thus, as suggested by Cutler et al, competition effects are indeed direct.

This brings us to the simulations of SHORTLIST. As a modeller, one has to

make a number of decisions as to which aspects one should pay attention to and to which not. Those which are neglected will hopefully turn out to be just implementational details. But, what is trivial and what is not, is difficult to decide on beforehand. What may look in the beginning as just a detail, may in the end turn out to be crucial. Returning to SHORTLIST, it is clear that much attention has been paid to getting the competition effects right, and it has to be acknowledged, SHORTLIST is one of the fist models which succeeds in it. But, one aspect already alluded to are the ubiquitous frequency effects. These are not just simple main effects of frequency, but they interact with competition effects. This may turn out to be a challenge for SHORTLIST. Another challenge is the translation from an activation pattern changing in time into reaction times, which are, in the end, the behavioral data on which our models are based. So far, SHORTLIST presents activation patterns, and we as *observer* can decide that the activation of item X at time T is higher than that of item Y. But, it will be much harder, if not insurmountable, if the machine has to make the right decision *at the right time*. Take as an example the comparison between *nemess* and *mestem*. The behavioral data show that *mess* is easier to detect in *nemess* than in *mestem*. In the SHORTLIST simulation, however, *mestem* is more highly activated than *nemess* at the critical C. One phoneme later, however, things are reversed, and three phonemes later, things are equal. What should we do, and even more important, what should the system do?

How to combine metrical segmentation and lexical competition?

There are different ways one may envisage the combination of a pre-lexical strategy like the MSS and a lexical competition procedure. One simple and straightforward conception is a two stage and hierarchical one. The speech stream is first segmented at the pre-lexical level and next, on the basis of the results of such pre-processing, the lexicon is accessed. Such a picture is attractive, both conceptually and empirically. It makes room for a number of answers to questions that have been investigated in the last couple of years. For example, the notion of a pre-lexical processing level offers a niche for linguistic differences - e.g., the presence or absence of syllabic effects in segment detection - that can be conceptualised independently of

the lexicon (note that we say 'conceptualised', not 'are'). This notion of a lexicon-independent level suits the results from infant studies. The prelexical procedure would, in the infant case, function as a bootstrapping strategy for building up a lexicon, and eventually, it may also be of help for the second language learner. What complicates this attractive picture is that a prelexical strategy like the MSS and lexical processes like competition and inhibition are not independent of each other. The MSS is not totally independent from the lexicon since the regularities speakers apparently exploit, are properties of the lexicon itself. Moreover, even considered at the early age, as a bootstrapping strategy for lexicon acquisition, the MSS must itself be the object of fine tuning to the relevant rhythmic properties of the native language. In other words, if we agree that in essence understanding language is a matter of understanding words, then there is no denying that properties of the language are defined over words. The bootstrapping procedure, however, may not be the same or work the same way with native and second language learners, especially when in the latter case we consider older subjects. For all that matter, one may envisage in adult second language learners that the lexicon is acquired on the basis of an analogy procedure. First a limited set of words is acquired. The language learning system may take this as its data basis and formulate a learning algorithm for lexical expansion based on the properties of the initial corpus. The next time round, a new and more powerful algorithm is formulated and serves as a pool of new and better hypotheses. On this picture, prelexical segmentation grows on lexical knowledge. This is not to say that the notion of a prelexical segmentation level has had its best time. The issue is not one that can be settled by results from simulation studies on their own. One may imagine that at a neurological and neuropsychological level, pre-lexical and lexical processes do indeed appear as rather different, even if behaviorally they cannot be unpacked.

To return to our opening remarks, then, whatever the conceptual coherence of linguistic notions, they are transported to the domain of human language processing and their operational edge is put to test. The problems of bringing the processing reality into focus should testify to their vigour and generative power. Achievements in second language acquisition or development in the native language are then not a matter of replacing a set of segments or a type of strategy with a new segment or

superimposing as level a new on an old one. It remains to be seen of course if in the long run, linguistic concepts are fit for the demands of dynamic explanations that first and second language acquisition confront us with. But we are just beginning to see what those demands are.

References

Bard, E. G. (1990). Competition, lateral inhibition, and frequency: comments on the chapters of Frauenfelder and Peeters, Marslen-Wilson and others. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing.* Cambridge, Mass: MIT Press.

Cutler, A., Norris, D & McQueen, J. (1994). Modelling lexical access from continuous speech input. Dokkyo University, Annual report of the International Center, 1994.

de Gelder, B., van Zon, M., & Vroomen, J. (in preparation). Non-native speakers go Dutch: evidence for a metrical segmentation strategy in speakers of French?

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighbourhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing.* Cambridge, Mass: MIT Press.

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

van Zon, M., & de Gelder, B. (1993). Perception of word boundaries by Dutch listeners. Proceedings of the Third European Conference on Speech Communication and Technology. Berlin, 689-692.

Vroomen, J., & de Gelder, B. (submitted). Metrical segmentation and lexical inhibition in spoken word recognition.

Vroomen, J., & de Gelder, B. (in preparation). Frequency effects and lexical inhibition in spoken word recognition.

Vroomen, J., van Zon, M., & de Gelder, B. (in preparation). Cues to speech segmentation: evidence from juncture misperceptions and word spotting.