6 Language-specific Processing: Does the Evidence Converge?

Anne Cutler
MRC Applied Psychology Unit, Cambridge, UK.

Language-specific processing is not a concept that psycholinguists leap to embrace. If all language processing were specific to the language being processed, Psycholinguistics would be a very different discipline; among other things, it would forfeit much of its interest in the eyes of its superordinate held, cognitive psychology. Cognitive psychologists study the structures and processes involved in human cognition, and therefore, when they study language processing, they are interested in human language processing in general rather than processing in the individual case. By extension, language-specific processing simply seems not to be an interesting phenomenon.

Nevertheless, there is an increasing body of evidence in favour of language-specificity in certain aspects of language processing. Dupoux's chapter, and the wider conference discussion to which this chapter is also addressed, have both reviewed and extended this evidence. The conclusion that emerges is an encouraging one: not only is language-specific processing an extremely interesting phenomenon, it also poses no problem to Psycholinguistics as a sub-discipline of cognitive psychology, because it can in fact be tractably interpreted within a language-universal framework.

The evidence that first put language-specific processing in the spotlight came mostly from experiments with a single technique, syllable-monitoring; studies in French, English, Spanish, and Catalan produced unexpectedly disparate results. The experiments measured listeners' response time to detect a match between a target CV (e.g. BA-) or CVC (e.g. BAL-) sequence and the initial sounds of a word. In French, RT was crucially determined by whether or not the target corresponded exactly to a syllable

of the word (Mehler, Dommergues, Frauenfelder, & Segui, 1981); thus RT to BA- was faster in *balance,* which has an open initial syllable, than in *balcon,* which has a closed initial syllable, while RT to BAL- was faster in *balcon* than in *balance.* In English, in contrast, there was no effect of syllabicity, but a strong effect of whether the target-bearing word began CVCV *{balance)* or CVCC *{balcony;* Cutler, Mehler, Norris, & Segui, 1986); this latter finding perhaps reflects the relative frequency of these patterns in the English vocabulary (Cutler, Norris, & Williams, 1987). The English result was partially replicated by Bradley, Sanchez-Casas, and Garcia-Albea (1993): they also found no effect of syllabicity, but also no effect of word structure. In Spanish, however, they found exactly the effect of syllabicity that Mehler et al. (1981) had observed in French; but this result in turn was not replicated by Sebastian-Galles, Dupoux, Segui, and Mehler (1992), who conducted further such experiments in Spanish and Catalan. In Spanish, there was only an effect of target size, whereas in Catalan there was a syllabicity effect which appeared only in words with non-initial stress, and a target size effect which appeared only in words with initial stress. Varying word stress in English, however, had no analogous effect—whatever the stress pattern, no syllabicity effect appeared (T. Mintz, personal communication; Rhodes-Morrison, 1992).

Given the psycholinguistic aim to study the human language processing system in general, it was somewhat disturbing to obtain such a variety of results from the same experimental paradigm in languages that are, historically, quite closely related. Some of the disparities—notably the disparities between two experiments on the same language—could be explained in terms of methodological artefacts. The odd study out is that of Bradley et al. (1993), which produced results on English that differed from those of Cutler et al. (1986), Mintz (personal communication), and Rhodes-Morrison (1992); and results on Spanish that differed from those of Sebastian et al. (1992). The way in which Bradley et al.'s study differed methodologically from the others is that it included foils, i.e. words containing a syllable that is highly similar but not identical to the target syllable (for instance, *badger* in a list with the target BAL-). When foils are included, subjects in syllable-monitoring experiments adopt a cautious response criterion (Norris & Cutler, 1988) and hence respond more slowly (indeed, the RTs reported by Bradley et al. are slower than the RTs in all the other experiments). As Dupoux (this volume) has demonstrated, response patterns are crucially dependent on subjects' response criterion; in particular, syllabicity effects are more likely when subjects respond slowly than when they respond rapidly.

The disparities between the results with different languages, however, could not be explained in such terms. Irrespective of experimental detail, no experiment found a syllabicity effect for English; but such effects did

appear, under appropriate conditions, for French, for Spanish, and for Catalan. Instead, the researchers invoked as explanation phonological differences between the four languages. Cutler et al. (1986) invoked rhythmic structure in explaining the French-English differences; whereas French has a basically syllabic rhythm, they argued, English is characterised by stress rhythm, involving especially the opposition of strong and weak syllables (i.e. syllables with full *vs* reduced vowels). Importantly, syllabic rhythm is also characteristic of Spanish and Catalan. Similarly, Sebastian et al. (1992) pointed to the difference in the size of the vowel inventory (large in French, small in Spanish) in accounting for the French-Spanish difference, and to the presence of reduced vowels in Catalan in explaining the Spanish-Catalan difference. According to their account, the larger the vocalic inventory, the more likely a syllabicity effect is to be observed. Putting these explanations together, the results converge on a consistent and plausible model. Firstly, it appears that an absolute prerequisite for the appearance of syllabically based responding in a syllable-monitoring task is that the language has syllabic rhythm. Given this, however, syllabic responding is most likely if the language has a large vowel inventory (French), somewhat less likely if the vowel inventory is intermediate (Catalan), and least likely if the vowel inventory is small (Spanish). In the latter case, syllabic responding will only appear when experimental conditions encourage subjects to adopt a cautious response criterion.

Subsequently, this series of studies was extended to a language that is *not* related to the four already discussed. Japanese does not have syllable rhythm, but it does have a clear and simple phonological structure, and very simple syllable structure. Exactly analogous experiments to those conducted in the four European languages produced no trace of syllabic effects with Japanese listeners (Otake, Hatano, Cutler, & Mehler, 1993). Instead, the Japanese listeners appeared to be basing their responses on mora units; the mora, a sub-syllabic unit which can be of the form V, CV, CCV, or syllabic coda, is the unit of rhythm in Japanese.

The balance of evidence therefore suggests that certain characteristics of a language's phonological repertoire exercise constraints on processing by native speakers of the language; in other words, some part of language processing is language-specific. One of the most important aspects of this series of studies is the fact that they included fully cross-linguistic investigations of listeners' processing: not only native-language processing, but also listeners' performance with foreign-language input. These cross-linguistic comparisons showed that if a syllabic response pattern appeared in a listener's native language, the listener could also produce that response pattern with input in a foreign language (for instance, French listeners produced a syllabic response pattern with English and with Japanese); but

no listener could produce a response pattern with foreign language input if that pattern was not characteristic of the native language (thus English listeners did not respond syllabically with French input or Spanish input, and neither English nor French listeners produced a mora-based response pattern with Japanese input). These findings made it quite clear that language-specific processing was a property of the listener, not of the input being presented to the system.

This aspect of the earlier results led Cutler, Mehler, Norris, and Segui (1989; 1992) to extend their investigations with syllable-monitoring to French-English bilinguals; interestingly, they found processing differences even in bilinguals whose command of both languages was, to all intents and purposes, equivalent and perfect. Specifically, a measure of language dominance predicted whether or not a syllabic response pattern would be observed with French-language materials: if on this measure the subjects were classified as French-dominant, they produced a syllabic response pattern; if they were classified as English-dominant, they did not. The same measure predicted whether or not a stress-based response pattern could be observed with English-language materials. Evidence from a number of paradigms suggests that English listeners' response patterns in speech recognition are stress-based: for instance, segmentation errors show a consistent pattern, whereby word boundaries are inserted before strong syllables, but overlooked before weak syllables (Cutler & Butterfield, 1992). Similarly, segmentation at strong syllables predicts the interference patterns observed in a task in which real words are detected in nonsense bisyllables (Cutler & Norris, 1988). Testing the French-English bilinguals on this latter task, Cutler et al. (1992) found that only subjects who were classified as English-dominant gave evidence of stress-based responding. Cutler et al. concluded that both syllabic and stress-based response patterns reflect language-specific segmentation procedures, of which only one may be available to a language user's processing system; in bilinguals, the available procedure would be determined by some (as yet unspecified) critical period or experience in language development. The choice, however, was determined by rhythmic characteristics of the language: syllabic rhythm encouraged a syllabic segmentation procedure, stress-based rhythm a stress-based segmentation procedure.

The suggestion that language-specificity occurs at the level of *segmentation processes* clearly offers a way in which the language-specific evidence might be integrated within a universal model of processing. The model tentatively proposed by Cutler et al. (1992) and Otake et al. (1993) assumes the existence of specific processes which act to postulate likely word boundaries in continuous speech input. The continuity of spoken language poses a significant problem for the perceiver, since language can only be understood by accessing lexical representations, and these

representations are discrete. Only rarely does speech provide reliable and robust cues to where one word ends and the next begins. Procedures that exist specifically to deal with this problem by exploiting language experience to segment speech with maximum efficiency are logically independent of perceptual units (see Norris & Cutler, 1985, and Cutler & Norris, 1988, for further arguments on this point). As Cutler et al. (1992) point out, there is a remarkable similarity in the syllabic effects shown by French listeners and the stress-based effects shown by English listeners, in that both patterns reflect the basic *rhythmic structure* of the language in question—syllabic rhythm in the case of French, stress rhythm in the case of English. The mora-based response pattern shown by Japanese listeners admits of exactly the same interpretation. This offers the opportunity for restitution of the universal dimension (and, incidentally, severe restriction of the possible range of language-specificity): the child is born with the capacity to develop aids to segmentation which exploit linguistic rhythm. That these procedures turn out to be language-specific derives from the fact that rhythmic structure differs across languages.

Since this proposal, if true, would appear to have some considerable importance for psycholinguistic theory, it is perhaps unfortunate that the majority of the body of research summarised earlier comes from studies with a single task: monitoring for CV and CVC sequences in spoken words. It would seem to be desirable to assemble converging evidence, not only from many experiments (this is surely already available) but from many experimental paradigms. To what extent is this possible?

In a sense, studies that are essentially monolingual can provide relevant evidence. For instance, the studies of stress-based processing in English, mentioned earlier, provide crucial evidence for the involvement of rhythmic factors in processing in English. Although such studies are obviously useful in the bilingual case, they do not allow fully cross-linguistic comparisons, since they cannot be replicated in, say, French. This is because French does not have English-like stress; only in other stress languages is replication of stress studies possible. Of course, it is possible to test alternative hypotheses monolingually; thus Cutler and Norris (1988) were able to show that a syllabic hypothesis would have predicted a completely different pattern of data in their word-spotting experiment. In addition, it is possible to construct similar monolingual experiments, testing language-specific hypotheses, within a particular experimental paradigm; thus word-spotting in French shows clear syllabic response patterns (J. Segui, personal communication).

Some exploratory work presented at the conference by Sebastian and by Altmann used a completely different technique: the presentation to listeners of time-compressed speech. Such a technique is extremely attractive because it appears to be quite language-independent—no

language-specific phonological constructs are involved. In principle, speech in any language can be presented with any degree of compression. Previous work with this methodology has indeed suggested that there may be language-specific effects in how compression affects intelligibility (Wingfield, Buttet, & Sandoval, 1979). Sebastian's and Altmann's preliminary findings suggest possible effects both at this level, and at the level of perceptual adaptation to speech compression within *vs* across languages. If such effects prove reliable, this technique would allow fully cross-linguistic investigations.

A different technique was used by Mack (personal communication; Mack, Tierney, & Boyle, 1990), who studied the recognition of natural and synthetically produced sentences by native speakers of English and by highly proficient bilinguals. The materials were presented without any noise-masking, time-compression or other distortion, in American English. Unsurprisingly, recognition by native speakers of American English was at a very high level indeed for both the natural and the synthetic speech; native speakers of another dialect of English (British) performed as well. Bilinguals, however, performed less well, despite the fact that Mack's bilingual subject population was exclusively composed of individuals who had been living and performing at a high academic level in an English-language environment for many years. Moreover, how well the bilinguals performed was apparently determined by their first language: bilinguals whose first language was German performed well (over 90% correct) with naturally produced speech, and less well (70% correct) with synthesised speech, but bilinguals whose first language was French scored only 73% correct for natural speech and 49% correct for synthesised speech—i.e. they performed significantly less well than the German-English bilinguals. Rhythmically, German resembles English to a far greater degree than French does, and it is clearly possible that Mack's studies have tapped into a difference in segmentation performance in a manner analogous to the syllable-monitoring studies. Thus Mack's results provide converging evidence with the Cutler et al. (1989; 1992) finding of unexpected processing limitations in bilingual speakers.

Yet another technique, and one that is particularly suited to the investigation of units of speech segmentation, is the illusory word detection task devised by Kolinsky and Morais (Kolinsky, 1992; Kolinsky & Morais, 1992). With this technique listeners are presented dichotically with competing signals and asked to judge whether one or other of them was a specified target word. The crucial experimental conditions occur when a subject's detection report would involve combination of parts of each of the two signals. For instance, with a target word *bijou,* a detection report from the simultaneous input of two nonwords *cojou* and *biton* would involve migration between the two input words at the syllabic level,

whereas a detection report from an input *kijou/boton* would involve migration of the initial phoneme. Preliminary investigations with French suggest that syllabic migrations are common in that language (Kolinsky, 1992; Kolinsky & Morais, 1992).

It may be said, then, that language-specific processing is already fairly well supported by converging evidence from diverse studies and tasks, and that the amount of available evidence is likely to increase. On one view, furthermore, the current evidence would seem to present a coherent picture, within which language-specificity does not conflict with the necessity of universality in a model of human language processing. Dupoux (this volume), however, offers a more radical interpretation, which seeks to restore universality at the perceptual unit level. The universal unit in terms of which speech is categorised is postulated to be the demisyllable. Demisyllables are combined at a later stage of processing into a second universal unit, the syllable. Language-specific differences are localised at the lexical level, in that it is only in some languages that the universal perceptual units also serve as the units of lexical access.

In postulating perceptual units that are not necessarily units of lexical access, Dupoux's proposal would seem not to have advantages of parsimony on its side. It would also seem to predict that the effects of response criterion on response patterns in the syllable-monitoring experiments ought to be constant across languages. Dupoux shows an impressive correlation in the French experiments between the size of the syllabic effect and the subjects' mean response times. However, post-hoc analysis of the three experiments reported by Cutler et al. (1986) with English listeners—experiments with English words, French words, and 'English' nonwords—does not support the constancy prediction. None of the three experiments shows a correlation between mean RT and syllabic effect; nor is there a correlation between mean RT and word effect (CV), as found by Dupoux for the French data. The correlation of mean RT with word effect (CVC), which Dupoux found to be insignificant for French listeners with French words, is also insignificant for English listeners with English words, but is significant for these listeners with both French words and 'English' nonwords; however, the two significant correlations are in the opposite direction to one another. It is unclear what one can make of this latter contradiction; what *is* clear, however, is that the results of these correlations for English listeners do not mimic the results produced by French listeners. If there is indeed universality across languages at the perceptual unit level, then its failure to manifest itself in the syllable-monitoring comparison of French and English requires independent explanation.

Although the detailed implications of Dupoux's proposal are not spelt out, his proposal is evidence that language-specific processing effects may

be explained within a universal model of language processing in more than one way. The rhythm-based segmentation proposal of Cutler et al. (1992) and Otake et al. (1993) also stands in need of fuller specification. Nevertheless, the existence of alternative proposals with the same basic aim is highly encouraging. The evidence is indeed converging: part of human language processing *is* language-specific, but in no way does this language-specificity serve to counter the essential universality of the human language processing system.

ACKNOWLEDGEMENT

Financial support is acknowledged from the Human Frontier Science Program.

REFERENCES

Bradley, D.C., Sanchez-Casas, R., &. Garda-Albea, J. (1993). The status of the syllable in the perception of Spanish and English. ***Language A Cognitive Processes, 8,*** **197-234.**

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. ***Journal of Memory and Language, 31,*** **218-236.**

Cutler, A., Mehler, J., Norris, D.G., ***A*** **Segui, J. (1986). The syllable's differing role in the segmentation of French and English.** ***Journal of Memory and Language, 25,*** **385-400.**

Cutler, A., Mehler, J., Norris, D.G., & Segui, J. (1989). Limits on bilingualism. ***Nature, 340,229-230.***

Cutler, A., Mehler, J., Norris, D.G., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. ***Cognitive Psychology, 24,*** **381-410.**

Cutler, A., & Norris, D.G. (1988). The role of strong syllables in segmentation for lexical access. ***Journal of Experimental Psychology: Human Perception A Performance, 14,*** **113-121.**

Cutler, A., Norris, D.G., & Williams, J.N. (1987). A note on the role of phonological expectations in speech segmentation. ***Journal of Memory and Language, 26,*** **480-487.**

Kolinsky, R. (1992). Conjunction errors as a tool for the study of perceptual processing. In J. Alegria, D. Holender, J. Morais, & M. Radeau (Eds.), ***Analytic approaches to human cognition,*** **pp. 133-149. Amsterdam: North Holland.**

Kolinsky, R., & Morais, J. (1992). Representations intermediates dans la reconnaissance de la parole: Apports de la technique de creation de mots illusoires. ***Actes da 19emes Joumees d'Etudes sur la Parole,*** **pp. 129-133, Brussels.**

Mack, M., Tierney, J., & Boyle, M.E.T. (1990). ***The intelligibility of natural and LPC-vocoded words and sentences presented to native and non-native speakers of English.*** **(Technical Report No. 869). Cambridge, MA: Lincoln Laboratory.**

Mehler, J., Dommergues, J.Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. ***Journal of Verbal Learning and Verbal Behaviour, 20,*** **298-305.**

Norris, D.G., & Cutler, A. (1985). Juncture detection. ***Linguistics, 23,*** **689-705.**

Norris, D.G., & Cutler, A. (1988). The relative accessibility of phonemes and syllables. ***Perception A Psychophysics, 43,*** **541-550.**

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. ***Journal of Memory and Language, 32,*** **358-378.**

Rhodes-Morrison, F. (1992). ***The role of sublexical units in speech segmentation.*** **Undergraduate project, Hatfield Polytechnic, UK.**

Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. ***Journal of Memory and Language, 31,*** **18-32.**

Wingfield, A., Buttet, J., & Sandoval, A.W. (1979). Intonation and intelligibility of time-compressed speedi. Supplementary report: English vs. French. ***Journal of Speech and Hearing Research, 22,*** **708-716.**