

## The processing of inflected forms

Charles Clifton, Jr.<sup>a</sup>, Anne Cutler<sup>b</sup>, James M. McQueen<sup>b</sup>,  
and Brit van Ooijen<sup>c</sup>

<sup>a</sup>Department of Psychology, University of Massachusetts, Amherst, MA 01003; <sup>b</sup>Max-Planck-Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands; <sup>c</sup>Laboratoire Sciences Cognitives et Psycholinguistique CNRS, Paris 75006, France, [cec@psych.umass.edu](mailto:cec@psych.umass.edu)  
[www-unix.oit.umass.edu/~cec](http://www-unix.oit.umass.edu/~cec)  
{[anne.cutler](mailto:anne.cutler@mpi.nl); [james.mcqueen](mailto:james.mcqueen@mpi.nl)}@mpi.nl  
[www.mpi.nl/world/persons/profession/{anne.html}; james.html](http://www.mpi.nl/world/persons/profession/{anne.html}; james.html)  
[brit@lscp.ehess.fr](mailto:brit@lscp.ehess.fr)

Abstract: Clahsen proposes two distinct processing routes, for regularly and irregularly inflected forms, respectively, and thus is apparently making a psychological claim. We argue that his position, which embodies a strictly linguistic perspective, does not constitute a psychological processing model.

Clahsen's argument is based on the inflectional system of German, because German allows comparisons not possible in English. This fact highlights yet again the problem that too much psycholinguistic theorizing relies exclusively on English, a language that is in many respects atypical (Cutler 1997). Clahsen claims that the German data motivate distinct processing routes for regularly and irregularly inflected forms. This appears to be a psychological claim. We argue, however, that his account, arising from a strictly linguistic perspective, does not constitute a psychological model at all.

Models of processing make behavioral as well as informational distinctions. It is not enough in a processing model simply to note that two distinct types of information may play roles in processing some material. A processing model must distinguish between, for example, recognition and production processes or, within recognition, between (modality-specific) access representations and (modality-independent) central representations. In fact the latter is a common distinction in psychological models of morphological processing (McQueen & Cutler 1998). No such distinction is made by Clahsen, however. It is thus impossible to say where his claims about the decomposition of regulars apply in the lexical access process - only centrally, or also at the access level? Does Clahsen rule out the possibility of full-form access representations for regular inflections, even if there are no such central representations?

To illustrate our argument, we report the results of an experiment in which listeners wrote to dictation ambiguous English word forms (McQueen et al. 1992; Van Ooijen et al. 1992). The experiment was based on an earlier study by Taft (1978), who showed that listeners preferred to write monomorphemic rather than inflected forms for ambiguous items such as [best]; note that prior studies (Bond 1973) have shown such items to be fully ambiguous between an inflected and a monomorphemic form (here, *based/baste*). We used Taft's task to compare regular and irregular inflections. The results, though clearly showing a processing difference between regularly and irregularly inflected forms, cannot easily be fitted into Clahsen's model.

The listeners were 32 native speakers of British English. They wrote 17 items involving a regular inflection and 19 items involving an irregular inflection, namely, regular: *billed/build* (.06), *paws/pause* (.06), *rays/raise* (.03), *sighed/side* (.03), *prints /prince*

(-44), *guessed/guest* (.13), *tents/tense* (.72), *passed/past* (.41), *paced/paste* (.34), *missed/mist* (.34), *knows/nose* (.09), *tied/tide* (.31), *guys/guise* (.19), *packed/pact* (.88), *based/baste* (.72), *days/daze* (.19), *stayed/staid* (.53); and irregular: *bred/bread* (.09), *won/one* (.19), *lain/lane* (.22), *blew/blue* (.04), *rode/road* (.31), *sought/sort* (.97), *knew/new* (.03), *caught/court* (.94), *read/red* (.13), *fought/fort* (.80), *seen/scene* (.81), *led/lead* (.25), *flew/flu* (.75), *heard/herd* (.72), *made/maid* (.47), *sent/scent* (.78), *thrown/trone* (.81), *feet/feat* (.84), *taught/taut* (.91). The numbers in parentheses are the proportions of inflected form responses to each item. The test also included 56 further items: homophones involving two monomorphemic forms (e.g., *beat/beet*), homophones involving function words (*would/wood*), and nonhomophones, both inflected and monomorphemic, in a single list recorded by a speaker of southern British English.

For the regular-inflection homophones (e.g., *based/baste*), listeners wrote the inflected form (*based*) on 32% of trials and the monomorphemic form (*baste*) on 68% of trials, a significant difference ( $z = 8.27, p < .001$ ). For the irregular-inflection homophones (e.g., *blew/blue*), listeners wrote the inflected form (*blew*) on 53% of trials and the monomorphemic form (*blue*) on 47% of trials, a difference that was not significant ( $z = -1.42$ ). There was therefore a qualitative difference between the regular- and irregular-inflection homophones. This is not the whole story, however. Another major determiner of listeners' choice was word frequency. Over all 36 items, the proportion of inflected choices correlated significantly with three different frequency measures (from a British English frequency count; Johansson & Hofland 1989): the inflected form's log frequency,  $r(35) = .34, p < .02$ ; the monomorphemic form's log frequency,  $r(35) = -.30, p < .04$ ; and the difference in these frequencies,  $r(35) = .48, p < .001$ .

Our items were in fact chosen so that there were subsets in which either the inflected or the monomorphemic form was higher in frequency. For the irregular-inflection homophones, listeners tended to write down whichever form was more frequent: The inflected form was chosen on 37% of trials when it was lower in frequency (e.g., *blew/blue*) and on 70% of trials when it was higher in frequency (e.g., *heard/herd*). However, for the regular-inflection homophones, there was a bias towards the monomorphemic form even when the inflected form was more frequent: The inflected form was chosen on only 24% of trials when it was lower in frequency (e.g., *billed/build*) and on only 40% of trials when it was higher in frequency (e.g., *based/baste*).

This striking difference between regular and irregular inflections was significant. In an analysis of covariance in which the log frequencies of both the inflected and the monomorphemic forms were used as covariates, there was a significant effect of regularity, with reliably more inflected choices with the irregular-inflection homophones than with the regular-inflection homophones:  $F(1,32) = 5.59, MSe = .0749, p < .03$ . This regularity effect was also significant in an analysis of covariance with the difference in log frequencies as covariate:  $F(1,33) = 5.35, MSe = .0378, p < .03$ .

These results support the hypothesis that there is a processing distinction between regularly and irregularly inflected forms. We believe that the lack of separate central representations for regular inflections underlies the distinction. With a homophone involving an irregular inflection, listeners have (aside from the frequency bias) a straight choice between two simple representations, but, with a homophone involving a regular inflection, the choice is between one simple representation (the monomorphemic form) and something more complex for the inflected form (whatever one's preferred account of how decomposed forms are represented).

Our point here, however, is that this interpretation already goes beyond the framework offered by Clahsen. We think that the effect entails central representations because (1) the correlation for the regular-inflection homophones alone between log inflected form frequency and proportion of inflected choices was not significant, and (2) other evidence indicates that spoken regularly in-

flexed forms have full-form access representations (Baayen et al., in preparation). However, because Clahsen does not distinguish between access and central representations, we are unable to offer him these data as support for his model.