

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

r the regular-inflection homophones (e.g., *based/baste*), lis-  
s wrote the inflected form (*based*) on 32% of trials and the  
omorphemic form (*baste*) on 68% of trials, a significant dif-  
ice ( $z = 8.27, p < .001$ ). For the irregular-inflection homo-  
ies (e.g., *blew/blue*), listeners wrote the inflected form (*blew*)  
3% of trials and the monomorphemic form (*blue*) on 47% of  
, a difference that was not significant ( $z = -1.42$ ). There was  
ore a qualitative difference between the regular- and irreg-  
nflexion homophones. This is not the whole story, however.  
her major determinant of listeners' choice was word fre-  
cy. Over all 36 items, the proportion of inflected choices  
lated significantly with three different frequency measures  
a British English frequency count; Johansson & Hofland  
r(35) = .34, p < .02; the  
omorphemic form's log frequency,  $r(35) = -.30, p < .04$ ; and  
ifference in these frequencies,  $r(35) = .48, p < .001$ .

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

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

ese results support the hypothesis that there is a processing  
ction between regularly and irregularly inflected forms. We  
ve that the lack of separate central representations for regu-  
nflexions underlies the distinction. With a homophone in-  
g an irregular inflection, listeners have (aside from the  
ency bias) a straight choice between two simple representa-  
but, with a homophone involving a regular inflection, the  
e is between one simple representation (the monomor-  
ic form) and something more complex for the inflected form  
ever one's preferred account of how decomposed forms are  
sented).

r point here, however, is that this interpretation already goes  
id the framework offered by Clahsen. We think that the ef-  
ntails central representations because (1) the correlation for  
egular-inflection homophones alone between log inflected  
frequency and proportion of inflected choices was not sig-  
nt, 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 of-  
fer him these data as support for his model.

## Investigating lexical entries and rules: A typological perspective

Greville G. Corbett

Surrey Morphology Group, Department of Linguistic and International  
Studies, University of Surrey, Guildford, Surrey, GU2 5XH, England.  
g.corbett@surrey.ac.uk www.surrey.ac.uk/LIS/SMG/gcorbett.html

**Abstract:** Research into entries and rules based on English is hampered  
by the fact that crucial factors coincide. Clahsen avoids this problem and  
demonstrates his claims by working on German. Seen against the back-  
ground of the immense variety of the world's languages, this successful  
move is potentially the first of many possible ones: Several languages of-  
fer promising configurations of the factors relevant to Clahsen, and others  
present new challenges.

Clahsen makes the point cogently that many investigations into  
the respective roles of entries and rules are vitiated by being based  
on a language in which regularity, overt suffixation, and high type  
frequency coincide (see sects. 2, 4.4, and 5.2). Clahsen and his col-  
laborators are able to tease apart the confounding factors by work-  
ing on German rather than English. By leaving the overcrowded  
section of the laboratory, as it were, to work on a quieter bench,  
they make considerable progress. If we take a typological per-  
spective, viewing the variety offered by the world's 6,000 or so lan-  
guages, then Clahsen has moved a relatively short distance: Ger-  
man and English are closely related, and, when compared to  
Mayali, Tsakhur, Yup'ik, or Japanese, they appear almost identical.  
There is thus a wonderfully large laboratory on offer, much of it  
virtually unused. The purpose of this commentary is therefore to  
suggest that Clahsen's successful move can be repeated many  
times, with the prospect both of further progress in the issue he  
addresses and of new challenges. This theme will be illustrated  
from one of the problems Clahsen investigates, namely, the cate-  
gory of number, although similar points could be made with rela-  
tion to the other problem, tense.

Clahsen is concerned largely with binary oppositions; for in-  
stance, German and English are limited to a singular-plural op-  
position of number. This is a common system, but many languages  
have three number values: singular, dual (for two), and plural  
(more than two); Yup'ik is an example. Some have an additional  
value, a trial (for three, as in Larike) or a paucal (for a small num-  
ber, as in Paamese). The largest systems have five number values  
(as in Sursurga). These large systems are found with pronouns  
rather than nouns, but duals on nouns are common. Given a sin-  
gular-dual-plural system, the question of regularity becomes more  
complex than in English or German, insofar as a lexical item may  
be regular in respect of one number opposition but irregular in re-  
spect of another (see, e.g., Priestly 1993 on Slovene, p. 401). These  
larger systems also bring with them interesting frequency effects:  
Singulars are normally more frequent than plurals, but the reverse  
holds for certain lexical items, such as *teeth*; when a dual is avail-  
able, it is typically the least frequent value, though for some lex-  
ical items it is the most frequent. (The existence of dual-dominant  
nouns is of relevance for the work on singular- and plural-dominant  
nouns by Baayen, Levelt, and Haveman, reported in Levelt  
et al. 1999, sect. 5.3.5)

German and English are also similar in that number is an oblig-  
atory category. If I say *It seems there's a dog in the park* when I  
know there are several, then I am being at best misleading. How-  
ever, in many languages, such as Japanese, an equivalent example  
with *dog* would be fully appropriate when more than one is in-  
volved; the plural would be used if the speaker wanted to draw

special attention to the quantity, but it is not obligatory. This difference has implications for frequency of use of the different number values, for the input for language acquisition, and in some languages for the degree of agreement among speakers on the form of the plural. While typically there is a shared form for singular and nonspecific or "general," a language may have a unique form for nouns that is outside the number system; this is found in Bayso (Corbett & Hayward 1987).

German clearly outdoes English in terms of the degrees of irregularity, and this is important for Clahsen. From a broader perspective, however, German nouns fall into a moderate number of (relatively) neat classes when compared to Nilotic languages, such as Shilluk. Gilley (1992, p. 81) quotes Kohlen, who following 30 years of study concluded that "A general rule for the formation of plurals in Shilluk cannot be given" (1933, p. 19). After detailed analysis, Gilley concurs. For many nouns one form is the base to which a suffix is added (whether to form the singular or the plural), whereas for some nouns both forms have a suffix. The real difficulties arise within the stem, which is typically of the shape consonant-vowel-consonant. There are some consonant alternations between singular and plural, and the vowel has four sources of variation, in terms of height, quality, tone, and length. It is as though the English plurals of the *foot* - *feet* and *mouse* - *mice* type made up a significant proportion of the lexicon (and with more possible types of alternation). Given the degree of unpredictable variation, Gilley (1992, p. 190) concludes that singulars and plurals in Shilluk cannot be derived from a single underlying representation and that two forms must be stored for each noun. Of course, it is difficult to prove a negative, but Gilley makes a good case. Yet Shilluk, with its remarkable degree of irregularity, provides some unconfirmed support for Clahsen's position (sect. 5.1). Gilley reports a suggestion that Shilluk children generalize the use of a suffix for the plural and then later learn the more complex forms (1992, pp. 130-31). Given the degree of irregularity, much greater than that of German, Shilluk suggests tantalizing questions regarding how far the balance between lexical entries and rules can be tipped in favour of one or other in particular languages.

In conclusion, Clahsen has shown again how beneficial it is to see English not as if it were Language but, rather, as just one (typologically rather odd) language. By moving the focus to German, he immediately clarifies the issues he wishes to address. There are many similar moves to be made, and, when we look at languages that are radically different from English and German, they can lead us to ask questions that English and German would not provoke. They offer new research space and new challenges. Unfortunately large parts of the laboratory are closing: According to Krauss (1992, p. 6), at the present rate, 50% of the world's 6,000 languages are likely to become extinct during the next 100 years.

## Chomsky's new clothes

Werner Deutsch and Oliver Müller

Department of Developmental Psychology, Institute for Psychology,  
Technical University Braunschweig, D-38106 Braunschweig, Germany;  
(w.deutsch; omueller)@tu-bs.de

**Abstract:** Clahsen's view on language is intimately linked with the Chomskian distinction between competence and performance. He uses performance to verify theoretical assumptions about the underlying structure of competence. Using mostly off-line tasks, he may fail to answer the question of how language is generated and perceived in natural situations.

In the old days of generative grammar, Chomsky introduced the distinction between the competence and performance of a human speaker/hearer. In Clahsen's account, competence is now labeled the language faculty consisting of innate, abstract, and modular organized linguistic knowledge, and performance is measured in a variety of tasks for language processing. The fit between the lin-

guistically defined genotype of language and the empirically obtained phenotype of language use in the area of morphology could not be better, could it?

First, the purpose of the empirical studies presented is to verify the proposed dual architecture of the language faculty. Why should data look like to have the slightest chance to revise the theory? Or is the theory a sort of biological ontology, resistant to falsification and modification?

Second, the variety of tasks Clahsen used in his endeavor to reflect and test the phenotypic appearance of two types of linguistic genes - the generative and the nongenerative ones - is impressive, indeed, but also restricted in a systematic way. Off-line tasks clearly predominate over on-line tasks. Do off-line tasks tap the same linguistic resources and procedures as on-line tasks do, or do they reflect "fallback procedures," as Butterworth (1983) has suggested? For example, how does access to the meaning of words affect their lexical and morphological processing?

Let us compare visual lexical decision for isolated German participles and the minimal requirements for accessing the meaning of such participles in a natural situation, characterized as follows: a continuous language/text flow, no expectation of nonwords. Furthermore, participles do not occur alone (except for ellipsis) but are embedded in constructions that use auxiliaries: the past participle (*Das Haus wurde gekauft*, "The house was bought") and the present/past perfect (*Sie hat/hatte ihn gesehen*, "She has/had seen him"). A hearer will accordingly expect a participle to occur when it is preceded by its accompanying auxiliary. (Of course, participles may be used in an adjectival function as in *Dies ist geschriebener Text*, "This is a written text." But Clahsen, Eisbeiss, and Sonnenstuhl [1997] excluded these forms in their frequency counts, presumably because they bear - in German, like English - inflectional suffixes.) The participles used by Clahsen et al. (1997) all bear the prefix *ge-*, which further signals the occurrence of a past participle. For the identification of participles (and their meaning) such as *ge-glaub-t*, "believe-d," or *schrieb-en*, "written," the phonemes or graphemes of the stem (*glaub* or *schrieb*) are accordingly sufficient and not all may be needed, depending on the specific uniqueness point. The suffixes *-t* or *-en* are redundant, in contrast to a lexical decision task where every bit of information is important in determining the yes/no response.

For a hearer in a natural situation it is only interesting to know which particular verb the speaker uses, because the grammatical information that a participle suffix signals is already specified in the earlier context. In the case of past participles, centering or access to meaning may lead to different processing, as in lexical decision. Clearly, language users are able to perform a wide range of tasks, but perhaps all these various performances are due to different competences or, preferably, to the joint activity of several processes whose contributions must be analyzed carefully, owing to one faculty.

Finally, what is the ultimate aim in studying language? Should it provide a realistic model of how a speaker generates an utterance such as "Harald has written a very interesting article" and how a hearer perceives, parses, and interprets such an utterance? Or is it about collecting positive evidence for a faculty, called *language*. That is the question!