

Bram Vandekerckhove, Emmanuel Keuleers, and Dominiek Sandra
University of Antwerp
bram.vandekerckhove@ua.ac.be, emmanuel.keuleers@ua.ac.be, dominiek.sandra@ua.ac.be

The role of phonological distance and relative support in the productivity of the Dutch simple past tense

Introduction

According to *dual-mechanism* accounts of inflectional morphology, regular inflection is governed by rules that operate over abstract symbols and is therefore fully productive, while irregular inflection depends on a database of stored word forms that allows limited productivity on the basis of similarity-based analogies. According to similarity-based *single-mechanism* models, both regular and irregular productivity depend on analogy-based processing of the target forms. In support of the dual-mechanism hypothesis, Prasada and Pinker (1993) (henceforth P&P) found that people's willingness to produce regular past tense forms for nonce verbs or to give these forms high ratings did not decrease with the nonce verbs' increasing *Phonological Distance* (PD) from existing regular verbs, while ratings and production numbers for irregularly inflected forms did decline with increasing PD of the nonce base forms from existing irregular verbs.

However, in their attempt to design stimuli that differed in their PD from existing irregular and regular verbs, P&P also caused another variable to shift, namely the relative frequencies of the morphological patterns among the closest phonological neighbors of the stimuli, or their *Relative Support* (RS): an increase in PD from existing irregular verbs is accompanied by a rise in RS for the regular inflection while increased PD from regular verbs actually results in a more balanced RS for regular and irregular patterns. This is exemplified by the average number of regular and irregular English verbs that rhyme with the nonce verbs of the different PD classes the authors created (P&P, p. 12).

This finding suggests that the results of P&P should be easy to replicate with a single-mechanism *Memory-Based Language Processing* model (Daelemans & Van den Bosch, 2005) of inflectional morphology. This has been shown by Keuleers and Sandra (submitted) (see also Eddington, 2000). The behavior of such a single-mechanism model is solely determined by the RS for the inflectional patterns among the words in the model's exemplar memory that are closest to the target form. Although the distance between the target word and the words in the lexical memory determines the relative influence of the different lexical items, the global distance between a word and its lexical neighbors itself does not influence the model's inflection choices.

This means, however, that the question whether PD from existing verbs has a *different* effect on regular and irregular productivity is still largely unresolved. We investigated both the effects of PD and RS on the productivity of regular and irregular patterns of the Dutch simple past tense, which like the English past tense system includes one large productive suffixation pattern for regular verbs, and a number of smaller 'gangs' of vowel-change irregulars with limited productivity.

The predictions of three different models were compared: (1) a *Partial-Blocking Dual-Mechanism* model (PBDM), which does not allow any involvement of stored regular items in the productivity of the regular tense but which does allow limited analogy-based productivity of irregular patterns that is able to (partially) block the application of the regular rule, (2) a *Fallback Rule Dual-Mechanism* model (FRDM), which allows both regular and irregular generalization on the basis of lexicon-based analogies, but which uses a symbolic rule for the regular past tense when the lexicon fails to provide a solution because PD becomes too large, and (3) a *Memory-Based Single-Mechanism* model (MBSM), in which both regular and irregular productivity are determined by similarity-based analogies between the target verb and the most similar verbs in the lexicon.

Table 1 summarizes the predictions of the three models. PBDM predicts that there should not be any effect of RS on the productivity of the regular past tense that cannot be accounted for by the number of rule-blocking irregular neighbors. FRDM and MBSM predict a positive effect of RS on the productivity of both the regular and the irregular past tense. In both dual-mechanism models, PD must have a negative effect on the productivity of the irregular inflection, since otherwise the symbolic rule would never be able to come into play: if PD of a word to the items in the lexicon had no effect on their influence, the rule mechanism would never be able to overcome lexical blocking, since the influence of the lexicon would always be equally strong. In PBDM, if PD has any effect at all on the productivity of the regular past tense, it should be a positive one, as increasing PD from the irregular items in the lexicon means less lexical blocking of the symbolic regular rule. It is not entirely clear what PD effect to expect for the regular inflection in FRDM. If there is a noticeable effect of PD on the productivity of the regular pattern it might be a positive one, since the rule operates under increasingly less lexical blocking when PD rises, unless of course rule-generated regular forms receive lower support from the rule mechanism than regular forms that are highly supported by the lexicon. MBSM does not in itself predict a negative effect of PD on the productivity of regular or irregular items, but if it took PD into account, this effect should be negative and equally large for regular and for irregular items when RS is held constant. Both dual-mechanism models predict interaction effects between RS and PD, since the influence of the lexicon should decrease when distance increases. MBSM does not predict such an interaction effect.

Method

Both the effects of PD and RS on the productivity of the Dutch simple past tense were investigated by having participants give acceptability ratings to past tense forms of nonce verbs that varied independently along those two dimensions for the regular

Table 1. Predictions of the three models under investigation concerning the effects of *Relative Support* (RS) and *Phonological Distance* (PD) on the productivity of the Dutch simple past tense.

	Main effect of RS		Main effect of PD		PD/RS interaction	
	Reg	Irr	Reg	Irr	Reg	Irr
PBDM	None	Pos	None/Pos	Neg	Yes	Yes
FRDM	Pos	Pos	None/Pos/Neg	Neg	Yes	Yes
MBSM	Pos	Pos	None (Neg)	None (Neg)	No	No

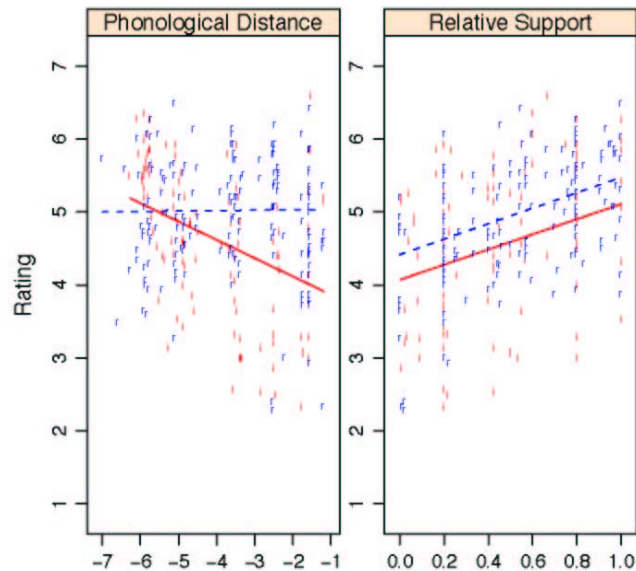
suffixation pattern and the 3 irregular vowel-change patterns with the highest type frequency. 144 monosyllabic nonce verbs were selected as stimuli for this experiment from a large pool of syllables whose phonological representations were assigned to past tense classes using the algorithms from *Tilburg Memory Based Learner* (TiMBL, Daelemans et al., 2008). The training set consisted of the phonological representations of monosyllabic verb stems with their past tense classes, extracted from the Dutch part of CELEX (Baayen et al., 1995). Stimuli were selected by crossing RS for each of the four patterns with mean PD to the nearest neighbors group. Participants were asked to rate the acceptability of simple past tense forms for each nonce verb on a scale from one to seven. 29 undergraduate students from the literature and linguistics department at the University of Antwerp, all native speakers of Dutch, took part in the experiment.

Results and discussion

Mixed effects models of covariance with Participant and Item as crossed random effects were fitted to the ratings for regular forms and irregular forms in stepwise regression analyses. The results (see fig. 1) show equally large positive effects of RS for regular ($\hat{\beta} = 1.066$, $t(2403) = 6.453$, $p = .0001$) and irregular verbs ($\hat{\beta} = 1.04$, $t(1735) = 3.658$, $p = .0001$). It does not seem to be the case that this effect of RS on regularly inflected items can be attributed to partial blocking by irregular neighbors, since the correlation between RS and the ratings ($M = 0.54$, $SD = 0.30$, $r(136) = 0.49$, $p = 7.15 \times 10^{-10}$) is significantly higher than that between the ratings and the number of irregular items among the phonological neighbors ($M = 1.29$, $SD = 0.63$, $r(136) = -0.35$, $p = 3.09 \times 10^{-5}$), $Z = -2.33$, $p = .0197$. This seems to rule out PBDM. PD has a significant effect on the ratings for the irregular nonce verbs, $\hat{\beta} = -0.25$, $t(1736) = -4.587$, $p = .0001$. However, there is no significant effect of PD on the ratings for the regular nonce verbs, $\hat{\beta} = 0.0064$, $t(2402) = 0.205$, $p = .857$. MBSM cannot account for this behavior if it does not take PD into account as an independent variable. FRDM does predict this effect of PD. However, this model ideally also predicts an interaction effect between RS and PD, of which we could find no evidence. This means that, although both FRDM and MBSM come close to explaining this pattern of results, the results actually are not straightforwardly explained by any of the models under investigation.

These results lead us to consider some other explanations. The decision which morphological pattern to choose for a given target word might for instance be determined by RS and functions similarly for both regular suffixation and irregular vowel-changes, while in the actual formation of the past tense itself, regular and irregular productivity are differently affected by PD. Another possibility is that participants were very sensitive to the large informative value of the past tense suffix in rating the past tenses. In an irregularly inflected verb, changing only one of the elements of the forms can have dramatic consequences for its interpretation as a past tense, since the past tense meaning is carried by the whole form. A regularly inflected Dutch verb, however, carries its past tense meaning exclusively on its suffix. This means that one can increase PD to the regular neighbors without diminishing the past tense meaning of the whole verb form.

Figure 1. Partial effects of *Phonological Distance* and *Relative Support* on the ratings for regularly (blue r's, blue broken line) and irregularly (red i's, red full line) inflected nonce verbs.



Conclusions

Although further experiments are necessary to explore all possibilities, these preliminary findings seem to suggest that lexical analogy on the basis of stored regularly inflected verbs plays a crucial role in regular productivity. A partial blocking account in which the application of the regular rule is dependent on the output strength of the memory component does not seem to be able to explain this pattern of results.

References

- Baayen, R. Harald, Piepenbrock, Richard, and Gulikers, Leon 1995. The CELEX lexical database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Daelemans, Walter and Van den Bosch, Antal 2005. *Memory-based language processing*. Cambridge: Cambridge University Press.
- Daelemans, Walter, Zavrel, Jakub, van der Sloot, Ko, and van den Bosch, Antal 2008. *TiMBL: Tilburg Memory Based Learner, version 6.1, Reference Guide*. ILK Research Group Technical Report Series no. 07-07, available from <http://ilk.uvt.nl/timbl/>.
- Eddington, David 2000. Analogy and the dual-route model of morphology. *Lingua* 110(4): 281-298.
- Keuleers, Emmanuel and Sandra Dominiek submitted. Similarity and productivity in the English past tense.
- Prasada, Sandeep and Pinker, Steven 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8(1): 1-56.