

**Do skilled readers exploit inflectional spelling cues that do not mirror pronunciation cues?
An eye-movement study of morpho-syntactic parsing in Dutch***

**Gerard Kempen, Andress Kooij & Theo van Leeuwen
Department of Psychology
Leiden University**

Introduction

The orthographies of Dutch and French, unlike those of English and German, frequently keep inflectional suffixes apart that have identical pronunciations. Examples are *aimer, aimez, aimé, aimée, aimés, aimées* in French, and *besteden, bestede, besteedde, besteedden* in Dutch. This feature is a well-known cause of serious spelling errors and mistakes in beginning and experienced writers because it renders word form selection contingent, not only upon sound and meaning, but also on syntax. It seems likely that this dependency on syntactic structure is not restricted to writing but has repercussions on reading as well. This invites the inference that morpho-syntactic analyzers in skilled readers have learned to pick up inflectional cues that only exist in the visual (orthographic) domain. The present study explores the validity of this hypothesis in the context of verb form spelling in Dutch.

(1) *Die baron die vorig jaar nog een vermogen had vergokt nu zijn laatste centen.*

That baron who last year still a fortune had gambles now his last pennies
“That baron who owned a fortune until last year is now gambling away his
last pennies.”

(2) *Die baron die vorig jaar nog een vermogen had verliest nu zijn laatste centen.*

... loses ...

(3) *Die baron die vorig jaar nog een vermogen had verspeelt nu zijn laatste centen.*

... gambles ...

(4) *Die baron die vorig jaar nog een vermogen had verkwist nu zijn laatste centen.*

... squanders ...

(5) *Die baron die vorig jaar nog een vermogen had verwedt nu zijn laatste centen.*

... wagers ...

*A slightly different version of this Long Abstract was published in *Abstracts of the Orthography Workshop “What spelling changes”*. The workshop took place at the Max Planck Institute for Psycholinguistics in Nijmegen on November 6-7, 1997.

Consider examples (1) through (5). Sentence (1) elicits a strong garden-path reaction due to the ambiguity of both verbs: *had* is categorially ambiguous between main verb and auxiliary; *vergokt* is inflectionally ambiguous between past participle and finite verb. (The *-t* suffix added to the stem *vergok* marks either the third-person singular present-tense form or the past participle.)

Most readers prefer to take *had* as auxiliary and *vergokt* as past participle, for several reasons: *had* has a much higher frequency as an auxiliary than as a main verb; *vergokt* as past participle fits in perfectly with this interpretation, which furthermore satisfies the powerful parsing principle known as Right Association (also called Late Closure or Recency; cf. Kempen, 1996). Sentence (2), however, has no—or a very weak—garden-path character, presumably because *verliest* is interpretable as a finite verb only. It differs from the past participle *verloren* in both spelling and pronunciation. Sentence (3) illustrates an intermediate case: *verspeelt* (whose meaning is similar to that of *vergokt*) is phonologically identical to, but orthographically distinct from the past participle *verspeeld*. (The *-d* suffix is pronounced /t/, due to final devoicing.)

Sentences (4) and (5), with main verbs *verkwist* and *verwedt* are similar to (1) and (3), respectively, but have a different morphological status. *verkwist* results from *verkwist-t* as a consequence of degemination of word-final consonants. This makes it unclear whether the final *-t* belongs to the stem or constitutes a suffix (inlike the *-t* in *vergokt*, which must be the suffix). In *verwedt*, the final *-t* plays an unambiguous suffixal role—like in *verspeelt*. The homophonous past participle *verwed* results from *verwed-d* by degemination, rendering the status of the final *-d* unclear. As a consequence, the relationship between *verwedt* and *verwed* differs from that between *verspeelt* and *verspeeld*: the members of the latter pair both have an explicit suffix; in the former pair, only one member is suffixed explicitly.

Using sentence materials as exemplified by (1)-(5), we have explored the question posed in the title of this paper. If readers and spellers have learned to exploit the orthographic difference between homophonic verb endings, sentences like (3) and (5) (with *verspeelt/verspeeld* or *verwedt/verwed*) are predicted to cause milder garden-path effects than sentences like (1) and (4) (with *vergokt* or *verkwist*) -- if the readers are garden-pathed at all. No garden-pathing is predicted for sentences of type (2).

Method

For each of the sentences types exemplified by (1)-(5), we prepared three additional variants. For instance, the set that includes sentence (1) is the following:

- I1 F C- *Die baron die vorig jaar nog een vermogen had vergokt nu zijn laatste centen.*
 I1 F C+ *Die baron die vorig jaar nog een vermogen had, vergokt nu zijn laatste centen.*
 I1 P C- *Die baron die vorig jaar nog een vermogen had vergokt spendeert nu zijn laatste centen.*
 I1 P C+ *Die baron die vorig jaar nog een vermogen had vergokt, spendeert nu zijn laatste centen.*

The symbols preceding the examples are explained in Table I. The C+ variants were introduced in order to obtain an estimate of the garden-path effect. We reasoned that a comma clearly marks the transition from the relative clause to the main clause, thereby removing the ambiguity (if any) created by a verb form. The P-variants enabled us to verify that the past participle interpretation is indeed preferred. (The P-versions contain a third verb that is unambiguously finite; in example (1), this is the verb *spenderen*, to spend.)

Table I. Experimental conditions of the experiment and their abbreviations. The symbols I1 and H1 denote verbs which undergo degemination in either the finite verb (F) or the past participle (P) forms, or both.

Comma	Verb Type		Finite Verb (F)	Past Participle (P)
Absent (C-)	Identical	I1	had verkwist	had verkwist
		I2	had vergokt	had vergokt
	Homophonous	H1	had verwedt	had verwed
		H2	had verspeelt	had verspeeld
	Different	D	had verliest	had verloren
	Present (C+)	Identical	I1	had, verkwist
I2			had, vergokt	had vergokt,
Homophonous		H1	had, verwedt	had verwed,
		H2	had, verspeelt	had verspeeld,
Different		D	had, verliest	had verloren,

We prepared 8 sentence quartets for each of the types I1, I2, H1 and H2, and 16 quartets for type D. Each of the participants (20 students of Leiden University who had been screened for their mastery of the verb spelling in Dutch) read one member of each these 48 quartets. The sentences were presented in quasi-random order, intermingled with 32 filler sentences of varying syntactic structure. At varying intervals, the participant had to answer a yes/no question about one of the recently presented sentences.

Results

The eye-movements during reading these sentences were registered by a Generation 5.5 Dual-Purkinje-Image eye-tracker. The data relevant for present purposes are the reading times (RT) for the 'ambiguous' (AMB) and 'disambiguating' (DIS) regions in the experimental sentences. These regions were defined as follows:

	AMB		DIS
Die baron die vorig jaar nog een vermogen / had, vergokt / nu zijn / laatste centen.			
	AMB		DIS
Die baron die vorig jaar nog een vermogen / had vergokt, / spendeert nu / zijn laatste centen.			

Table II. Estimates of the garden-path effect in the F- and P-variants of the five sentence types. The numbers are difference scores (milliseconds) computed by subtracting C+ RTs from C- RTs.

Region	Sentence type		Total Gaze		First Gaze	
	F	P	F	P	F	P
AMB	I1	VERGOKT/VERGOKT	123	36	33	11
	I2	VERKWIST/VERKWIST	168	16	41	23
	H1	VERWEDT/VERWED	115	47	29	4
	H2	VERSPEELT/VERSPEELD	224	35	61	1
	D	VERLIEST/VERLOREN	67	12	24	13
DIS	I1	VERGOKT/VERGOKT	128	59	42	17
	I2	VERKWIST/VERKWIST	108	20	64	14
	H1	VERWEDT/VERWED	86	92	-32	62
	H2	VERSPEELT/VERSPEELD	86	50	5	21
	D	VERLIEST/VERLOREN	21	18	-7	16

For each of the five sentence types, and for the F- as well as the P-versions, we obtained an estimate of the garden-path effect by subtracting the RTs in the 'easy' C+ condition from the 'difficult' C- condition. Table II presents the resulting difference scores for both regions, divided into 'First Gaze' scores and 'Total Gaze' scores. The First Gaze RT for a region starts at the first fixation within that region and stops when a position outside that region is fixated. That is, regressions into that region are not taken into account. The Total Gaze RT does include regressions and measures the total time spent in the region. (For details, see Konieczny et al., 1995.)

Leaving details of the statistical analyses aside (but see Kooij, 1997) we point out three remarkable outcomes. To begin with, the readers are hardly garden-pathed by the P-variants of the I1- and I2-sentences,

whereas the F-versions of these sentence types cause a great deal of trouble. This pattern is already already discernible in the First Gaze scores. On the other hand, D-type sentences hardly cause a garden-path effect. Both observations confirm our a priori expectations. Secondly, H2-type sentences in general behave very much like their I1- and I2-type counterparts. This phenomenon entails a negative answer to the title question: the orthographic but homophonic distinction between verb forms like *verspeelt* and *verspeeld* does not at all facilitate morpho-syntactic analysis. In the ambiguous region, the F-versions of this sentence type are even responsible for the largest garden-path effect of all (224 ms). Thirdly, type-H1 sentences cause a relatively mild garden-path effect. In the DISambiguating region there is no garden-oath effect at all! It follows that H1 verb forms like *verwedt* and *verwed* are indeed intermediate between fully identical F- and P-forms (*vergokt/vergokt*) and those which differ both in sound and image (*verliest/verloren*).

Discussion

The most salient finding of this study is the split between two types of homophonic but non-homographic inflectional cues. A verb form like *verwedt* is easily recognized as a finite verb, implying a positive answer to the title question. On the other hand, readers are troubled by forms such as *verspeelt*, which entails a negative answer. This split appears reliable: we have observed it in two pilot studies with a self-paced reading task (Evenblij, 1995; Kooij, 1996). However, it was not observed by Van Heuven (1978, 1991; see also Tismeer, 1994), who arrived at an overall negative answer to the title question: visible but inaudible spelling cues do not guide the parsing process, neither in type-H1 nor in type-H2 verb forms. On the other hand, Brysbaert (1996) found that readers quickly recognize verb forms like *wiedden* (of *wieden*, to weed) and *roestten* (of *roesten*, to rust) as past-tense forms, despite the homophony with the present-tense forms (and with the infinitives).

A possible explanation for the discrepancy between Van Heuven's data on one hand and Brysbaert's and ours on the other may have to do with task differences. In Van Heuven's experiments, the subjects had to read the sentences aloud. Moreover, the sentences were presented in capital letters without any punctuation. This may have rendered audible cues much more salient than in normal silent reading, causing the parsing process to rely on them rather than on the—somewhat degraded—visible cues. (Brysbaert used a self-paced silent reading task.)

We conclude by suggesting an explanation for the H1 versus H2 split. It capitalizes on the fact that spelling patterns such as *-iedd-*, *-stt-*, and word-final *-dt* only occur in polymorphemic words. For instance, *-stt-* indicates past-tense formation (*roestten*) or composition (*feesttent*, party tent). Such spelling patterns perhaps trigger morpho-syntactic analyzers more directly than spelling patterns that occur in mono- as well as in polymorphemic words. (The latter is true of the word-final patterns in type-H2 forms.) We intend to follow up this suggestion in future eye-tracking studies.

References

- Brysbaert, M. (1996). *Sentence reading: Do we make use of non-audible cues?* Manuscript, Catholic University of Louvain.
- Evenblij, J. (1995). *The utilization of d/t verb suffixes during reading and parsing Dutch sentences*. M.A. Thesis, Department of Psychology, Leiden University.
- Heuven, V. van (1978). *Spelling en lezen: hoe tragisch zijn de werkwoordsvormen?* Assen: Van Gorcum.
- Heuven, V. van (1991). Invloeden van spelling op het lezen. In: Bennis, H. , Neijt, A. & Santen, A. van (Red.), *De groene spelling*. Amsterdam: Bakker.
- Kempen, G. (1995). Computational models of syntactic processing in human language comprehension. In: Dijkstra, T. & De snedt, K. (Eds.), *Computational psycholinguistics: symbolic and subsymbolic models of language processing*. London: Taylor & Francis.
- Konieczny, L., Hemforth, B., Scheepers, C. & Strube, G. (1995). PP-attachment in German: Results from eye movement studies. In: Findlay, J.M., Walker, R., & Kentridge, R. (Eds.), *Eye movement research: Mechanisms, processes and applications*. Amsterdam: Elsevier.
- Kooij, A. (1996). *Het gebruik van de werkwoordelijke d/t uitgangen tijdens lezen*. Internal Report, Department of Psychology, Leiden University.
- Kooij, A. (1997). *Helpt de werkwoordspelling de lezer bij het herkennen van zinsstructuur? Een oogbewegingsstudie*. M.A. Thesis, Department of Psychology, Leiden University.
- Tismeer, L. (1984). *Effecten van verklankbaarheid van geschreven verschillen tussen werkwoordsuitgangen bij hardop- en stillezen*. M.A. Thesis, Linguistics Department, Leiden University.