The perception of word stress in existing and non-existing Dutch words by native speakers and second language learners*

Johanneke Caspers
Leiden University Centre for Linguistics

1. Introduction

Languages differ with regard to their prosodic structure. This means that speech phenomena involving melody, duration and loudness may be organized in different ways in different languages. The question central to the current investigation is whether the prosodic structure of the first language (L1) influences the perception of word stress in Dutch as a second language (DSL). The prosodic make-up at the word level may differ rather fundamentally between languages. For example, Chinese has lexical tone, Polish has fixed prefinal word stress, French probably does not have stress at the word level and Dutch has variable word stress. It may be the case that the prosodic structure of a first language is important when acquiring a second language, but this field of research has received only limited attention over the past decades (Chun 2002). Recently however, interest in prosodic aspects of second language acquisition has been growing (see e.g. Trofimovic & Baker 2006, Munro & Derwing 2006). In this paper we investigate the influence of differences in L1 prosodic structure on the perception of word stress in Dutch.

In earlier research into the production of word stress in Dutch as a second language, a group of French and Chinese DSL-speakers and a control group of native speakers read aloud materials containing both simplex and morphologically complex words with regular and irregular stress patterns. The similarities between the results of the French and the Chinese DSL-speakers strongly outweighed the differences. It seemed that both groups had successfully acquired the basic principles of the Dutch stress system. The vast majority of erroneous stress assignments could be interpreted as resulting from the overgeneralization of Dutch phonological and morphological stress rules and could not be attributed to influence of the prosodic structure of the L1 (Caspers & van Santen 2006). This means that no evidence for prosodic transfer was found with respect to word stress production in DSL.

Linguistics in the Netherlands 2009, **25–38**. DOI 10.1075/avt.26.04cas
ISSN 0929–7332 / E-ISSN 1569–9919

Recent investigations of the perception of word stress in a second language show conflicting results. Altmann's (2006) investigation of stress perception in English as an L2 leads to the conclusion that a predictable stress position in the L1 (e.g., French) is problematic, while absence of word stress in the L1 (e.g., Chinese)[1] does not lead to problems with the perception of English word stress. On the other hand, French as well as Chinese learners of Polish displayed problems with perceiving stress position in nonsense words (Kijak 2009).

2. Research question and approach

The limited amount of available research suggests that the *production* of stress in Dutch as a second language is not influenced by prosodic transfer from the L1. The *perception* of word stress in Dutch by native speakers of languages with different word prosodic systems has not been experimentally investigated, but research on other L2s suggests that prosodic transfer may be relevant to stress perception. Therefore the following research question was formulated: does the prosodic structure of the L1 influence the perception of word stress in Dutch as a second language? To answer this question we presented groups of late DSL-learners with prosodically different mother tongues with Dutch words. The following variables were included in the materials: existing versus nonsense words, the presence versus absence of a pitch accent on the target word, the predictability of the stress position (regular versus irregular), the word length (two versus three syllables) and the stress location within the target word (on the initial, final or prefinal syllable).

As a rule, experimental studies of word stress are restricted to nonsense words (for an overview see Guion 2005, Altmann 2006), under the assumption that testing the perception of real words does not reveal phonological processes, since the stress positions could be lexically stored. However, we maintain that real words should be investigated as well, since this assumption may be incorrect, and therefore both existing and non-existing words were included in the current experiment, keeping their characteristics as comparable as possible.

A further point that has not received much attention in stress research so far is the effect of pitch accent. As far as we know, all perception studies work with accented target words (i.e., the stressed syllable is marked by a conspicuous pitch change, in addition to a longer duration and greater loudness), while word stress is present in non-accented words as well, albeit instantiated by fewer phonetic cues (only duration and loudness, cf. Van Heuven 2001). Therefore the presence versus absence of a pitch accent was a factor in our design. We presented the target words in a fixed carrier sentence:

Wil je nog een keer xxx zeggen
Would you again one time xxx say
'Would you say xxx again'

Either the target word itself ('xxx') or another word ('nog') in the carrier sentence was accented.

Half of the target words have a stress position that can be predicted on the basis of simple phonological rules, the other half have irregular stress patterns, since earlier work on stress production revealed a clear influence of this factor (cf. Caspers & van Santen 2006). Word length and stress position within the word were systematically varied. A full list of target words can be found in Appendix 1.

The task of the subjects was to listen to the target word and circle on an answer sheet the syllable that was perceived as stressed. For the native subjects we expected good overall performance. For the DSL-speakers in general we expected more difficulty in perceiving the location of word stress, especially for those subjects that have a fixed stress position in their L1. Furthermore, we expected perception of stress to be more difficult in the minus-accent condition.

3. Method

Stimulus materials

All sentences were read aloud by a trained speaker and stored on computer disk. The pitch accent locations and stress positions in the resulting materials were checked by three phonetically trained judges, using PRAAT (Boersma & Weenink 2006). After verification and segmentation, the digital recordings were placed in audio files. We presented the 192 stimuli in four separate sets, ordered from more difficult to more easy. A pilot test had revealed that detecting the stress position in unaccented target words was much harder than in accented target words. Therefore the nonsense words without a pitch accent were presented first, followed by the real words without a pitch accent, then the nonsense words with a pitch accent and in conclusion the accented real words. Within these four sets the order of presentation was random.

Subjects

For our experiment we needed groups of DSL-speakers with prosodically different L1s. In addition to subjects with French or a tone language such as Mandarin Chinese as L1, we selected DSL-speakers with a fixed stress position in their first language. In total we tested 51 DSL-subjects; they can be divided into the following five prosodic groups according to their L1:

1. French, fixed phrase-final stress (N = 18)
2. Chinese, lexical tone (N = 11)
3. Polish, fixed stress on prefinal syllable (N = 9)
4. Finnish and Hungarian, fixed stress on initial syllable (N = 5)
5. diverse languages (N = 8)[2]

All subjects had followed intensive Dutch language courses at university level for at least six months. Information about self-reported proficiency levels in Dutch is included in Appendix 2. The control group contained 20 native speakers of Dutch.

Procedure

The subjects listened to the stimuli at a comfortable loudness over Quad ESL-63 electrostatic loudspeakers, while seated in a sound treated lecture room. The sessions started with the distribution of answer booklets. On the first page a short questionnaire had to be filled in (year and place of birth, mother tongue, L3, L4, proficiency in Dutch, etc.). Then the non-native subjects were presented with an alphabetical list of the real Dutch target words; they were asked to indicate for each word whether they were familiar with it or not. Subjects were then given the opportunity to read the instructions; they were requested to decide for each stimulus where the stress was located and circle this most prominent syllable on their answer sheets. Each of the four parts of the experiment was preceded by a number of practice items.

4. Results

A Repeated Measures Analysis of Variance on the percentage of correct responses with type of L1 as the between-subjects factor and word type (real or nonsense), pitch accent (present or absent) and regularity of the stress position (regular or irregular) as the within-subjects factors showed main effects of type of L1 ($F_{5,65} = 3.096$, $p < .05$), presence versus absence of a pitch accent ($F_{1,65} = 219.666$, $p < .001$), word type ($F_{1,65} = 5.595$, $p < .05$) and regularity of the stress position ($F_{1,65} = 13.102$, $p < .005$). There are significant interactions between type of L1 and accent ($F_{5,65} = 2.371$, $p < .05$), between type of L1 and regularity ($F_{5,65} = 3.899$, $p < .005$), between word type and accent ($F_{1,65} = 13.224$, $p < .001$) and between regularity and accent ($F_{1,65} = 15.189$, $p < .001$); all other two- and three-way interactions are insignificant.

Figure 1 presents the mean percentage correct perception of stress location for the different groups of subjects, in descending order. What is striking is that there is a division between, on the one hand, the results of the native speakers, the Chinese and French DSL-speakers as well as the group with diverse L1s, and, on the other hand, the results of the Polish and Finno-Ugric DSL-speakers. In spite of the significant main effect of type of L1, a posthoc SNK analysis does not reveal significant differences between all relevant groups of subjects. This is probably due to the relatively small number of subjects per group and the high variance within groups. Collapsing the data of the Polish and Finno-Ugric speakers, however, does lead to a significant difference between this new group and each of the other four groups in a posthoc test. These results suggest that a mother tongue with a fixed word stress system causes more difficulty in perceiving stress in a second language with a variable system then an L1 without fixed word stress (i.e., French or Chinese). An unexpected result was the relatively poor performance of the group of native speakers (80% correct).

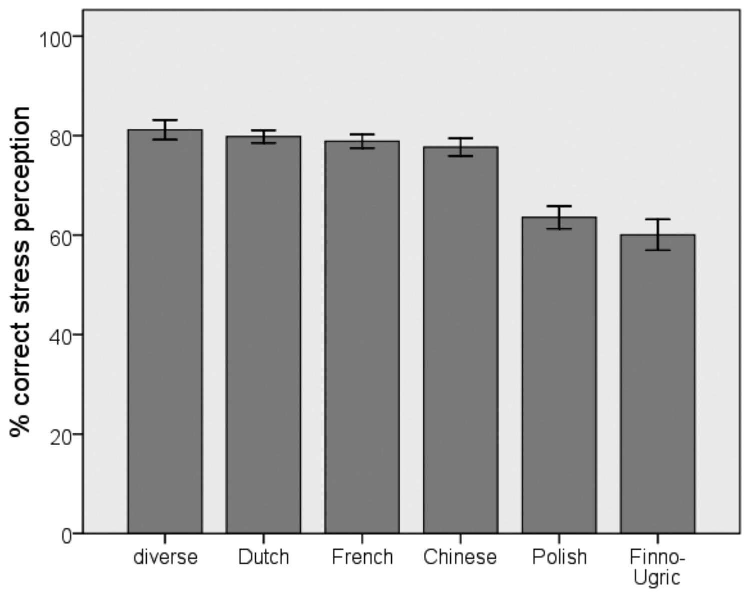

Figure 1. Percentage of correct stress perception for the different types of L1 (error bars: 95% confidence interval).

Figure 2 presents the effect of the presence versus absence of a pitch accent on the percentages of correct stress perception per L1 type. The figure clearly shows the large facilitating effect of pitch accent: when the target word is not accented, the percentage correct drops considerably. This effect is distinctly smaller for the Polish DSL-speakers (hence the significant interaction between type of L1 and presence of a pitch accent).

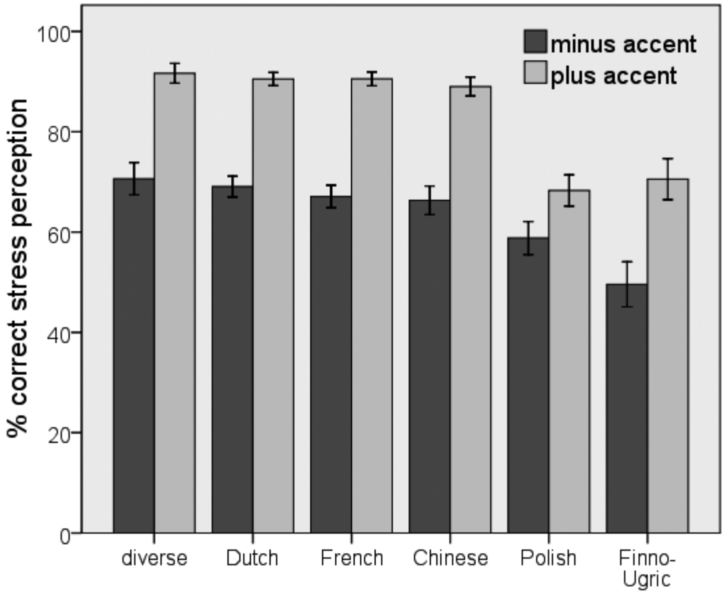

Figure 2. Percentage of correct stress perception for the different types of L1, broken down by the presence versus absence of a pitch accent on the target word.

The effect of word type — existing words versus phonologically possible but non-existent Dutch words — is relatively small. Overall the position of word stress in the real Dutch words is a little easier to perceive than in the nonsense words (see Table 1). There is no clear relationship between the number of existing words that were marked as known by the DSL-speakers before the perception experiment started and the performance on these words (cf. the second and third column in Table 1). Especially the Polish and Finno-Ugric speakers seem to have stored many of the words in their lexicon, but without (correct) stress information. The

Table 1. Percentage of known words and correct stress perception for the real words, plus the percentage of correct stress perception for the nonsense words.

	real words		**nonsense words**
L1	**known**	**correct stress perception**	**correct stress perception**
Dutch	–	80%	79%
diverse	75%	83%	79%
French	79%	80%	77%
Chinese	59%	78%	77%
Polish	86%	65%	61%
Finno-Ugric	82%	60%	60%

unknown words receive more correct responses than the known words (80% versus 74% respectively).[3]

The effect of regularity of the stress location, broken down by subject group, is presented in Figure 3. For the French and Finno-Ugric DSL-speakers regular word stress seems easier to perceive than irregular word stress.

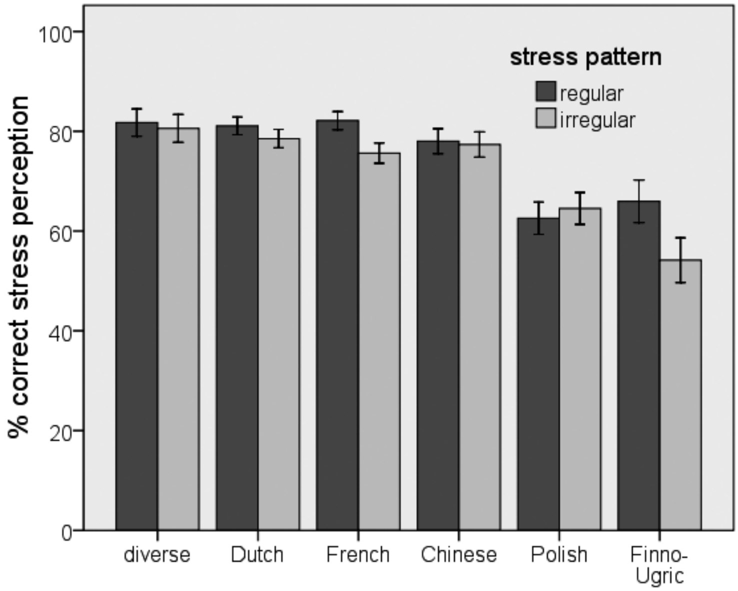

Figure 3. Percentage of correct stress perception for the different types of L1, broken down by regular versus irregular stress patterns.

The effect of regularity as well as word type is larger in the condition without a pitch accent than in the plus accent condition, hence the significant interactions between the factors word type and accent and between regularity and accent.

Finally, Figure 4 presents the effect of vowel length. This factor was not part of the original design, but posthoc inspection of the target words revealed that short and long stressed vowels are fairly evenly distributed over the different conditions. The effect appears to be rather large: a Repeated Measures ANOVA on the percentage correct with type of L1 as the between-subjects factor and vowel length (short versus long) as the within-subjects factor shows a main effect of vowel length ($F_{1,65} = 79.578$, $p < .001$); there is no interaction with L1 type ($F_{1,65} = 1.979$, ins.).

The fact that vowel length is not contrastive but used as a cue for word stress in Polish could provide an explanation for the effect found for the Polish DSL-speakers (Peperkamp & Dupoux 2002). However, Finnish and Hungarian do have contrastive vowel length, which makes the effect of vowel length on stress perception for this group of DSL-speakers difficult to understand. Also, French and

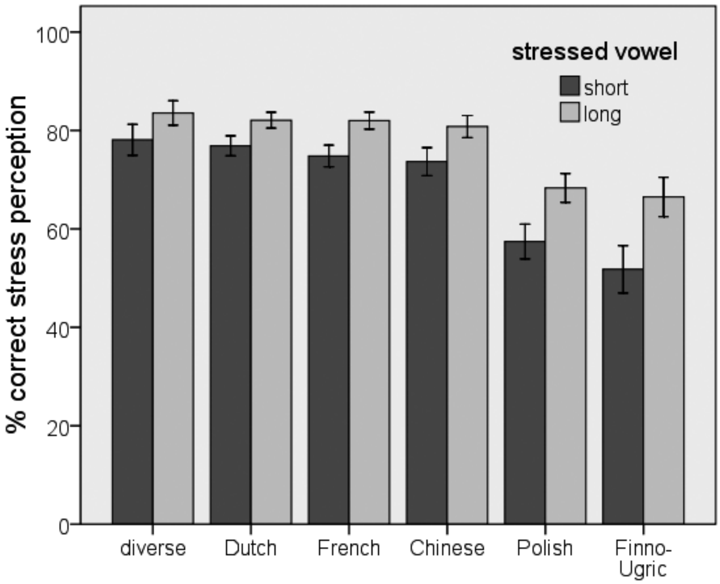

Figure 4. Percentage of correct stress perception for the different types of L1, broken down by vowel length of the stressed syllable.

Chinese, like Polish, do not have contrastive vowel length, while the performance of these speakers is comparable to that of the native speakers. There seems to be an overall facilitating effect of vowel length on the perception of word stress, even for the native speakers.

5. Conclusions and discussion

DSL-speakers with Polish as L1 perform worse (a mean of 64% correct responses) than speakers with other types of L1, except for the group of Finnish and Hungarian speakers of Dutch, which shows the lowest percentage of correct perception of stress position: 60%. The DSL-speakers with French or Chinese as L1 do much better (79% and 78% correct respectively); their results are almost as good as those of the group of native speakers. These findings indicate that a fixed stress position in the L1 interferes with stress perception in DSL, but absence of word stress does not: there is no significant difference in percentage of correct stress perception between the native speakers and the DSL-speakers with French or Chinese as L1. This is consistent with the results found earlier for Dutch stress production (Caspers & van Santen 2006), but the results are not in accordance with either the data found by Altmann (2006) or by Kijak (2009), see Section 1. A possible explanation for these inconsistencies is the L2 that was investigated — Dutch versus

English and Polish, respectively — but further research is necessary to be able to identify the precise cause for the conflicting findings.

The performance of the native speakers (a mean score of 80% correct) was lower than expected. We anticipated that the native speakers would outperform the non-native speakers; instead, they perform as good as the DSL-speakers with French or a tone-language as L1. Closer inspection of the data revealed substantial differences between individual subjects: some native speakers perceive stress position perfectly, others score near random, and the remainder of the subjects have a score somewhere between these two extremes.[4] This pattern can be seen in the other groups of subjects as well. It seems that people differ in the ability to perceive word stress position, as measured in this type of experiment.

There was a large facilitating effect of the presence of a pitch accent on the target word on the perception of the stress location (ca. 20% higher scores, cf. Figure 2). This effect was comparable for all groups of subjects, except for the Polish DSL-speakers: for these subjects the facilitating effect was much smaller (less than 10%). At present we do not have a satisfactory explanation for this finding.

The effect of word type (existing versus non-existing Dutch words) was small. In combination with the fact that the stress position in the unknown existing words was perceived better than in the known existing words by the DSL-speakers, this means that our data do not support the view that investigating stress perception should be done on nonsense words exclusively. The assumption that words are stored and can be retrieved with their — correct — stress position is not supported by the present data, since the native speakers do not perform better on the real words than on the nonsense words. We can safely assume that all existing words used in the experiment were known to the native speakers.

Further analyses of the influence of the nested factors target word length and stress position are needed, in combination with L1 influence.

One issue that must be considered, both with regard to the present study and to those of Altmann (2006) and Kijak (2009), is the possibility that the task performed by the subjects does not tap into a linguistically relevant process. French late learners of Spanish have much difficulty in performing on-line tasks involving phonological representations (they are 'stress-deaf'), whereas they are sensitive to the suprasegmental correlates of stress (Dupoux, Sebastián-Gallés, Navarrete & Peperkamp 2008). In principle this means that the results presented above may not give an adequate picture of the perception of word stress in running speech. However, the data do present a clear influence of the prosodic make-up of the L1: subjects with fixed (initial or prefinal) word stress in their L1 have much more difficulty in performing the off-line task of indicating the stress position of aurally presented Dutch words than subjects with no stress at the word level in their native language.

Notes

* The experiment was prepared and run in a research seminar under the guidance of the author and Ariane van Santen (LUCL); Olena Andrukh, Klaudia Bartków, Monique Bisschop, Diane Heuvelman, Martina Hoffmann, Sueandra Hollemans, Merlijn de Jager, Sven Klippel, Fleur Leijen, Marco Orru, Marja Oudega, Marlies Paalvast, Juliette Sandberg, Maartje Schulpen, Urszula Sniecinska, Karen Vegt, Dorina Veldhuis, Johanna Weeda and Manouk van Zetten participated in the research seminar as students. Laurent Rasier is thanked for running part of the experiment in Belgium. I am very grateful to Ariane van Santen for her comments on a draft version of this paper. Two anonymous reviewers are thanked for their very helpful comments. I am greatly indebted to Vincent van Heuven for his help in improving the statistical analysis of the data.

1. Note that it is controversial to consider French to have word final stress (Beyssade, Delais-Roussarie, Doetjes, Marandin & Rialland 2004); likewise, it is not undisputed to say that Chinese does not have word stress (Duanmu 2000).

2. None of these languages had phrase-level stress, lexical tone or fixed lexical stress.

3. Broken down by L1 the percentages correct responses for unknown versus known target words are as follows: diverse 84% vs. 82%, French 84% vs. 79%, Chinese 79% vs. 78%, Polish 73% vs. 64% and Finno-Ugric 65% vs. 59%.

4. A possible explanation for the relatively low scores of the control group is the fact that in most stress perception experiments reported on in the literature, subjects received a short training in stress perception, while in the present experiment the subjects did start each part of the experiment with a number of practice items, but they did not receive feedback on their responses to these items.

References

Altmann, H. 2006. The perception and production of second language stress: A cross-linguistic experimental study. PhD diss., University of Delaware.

Beyssade, C., E. Delais-Roussarie, J. Doetjes, J. Marandin & A. Rialland. 2004. "Prosodic, syntactic and pragmatic aspects of information structure, An introduction". *Handbook of French semantics* ed. by F. Corblin and H. de Swart, 455–475. Stanford, CA: CSLI Publications.

Boersma, P. & D. Weenink. 2006. Praat: Doing phonetics by computer (version 4.4.24). <http://www.praat.org>.

Caspers, J. & A. van Santen. 2006. "Nederlands uit Franse en Chinese mond: Invloed van T1 op de plaatsing van klemtoon in Nederlands als tweede taal?". *Nederlandse Taalkunde* 11, 289–318.

Chun, D.M. 2002. *Discourse intonation in L2: From theory and research to practice*. Amsterdam & Philadelphia: John Benjamins.

Duanmu, S. 2000. *The phonology of standard Chinese*. Oxford: Oxford University Press.

Dupoux, E., N. Sebastián-Gallés, E. Navarrete & S. Peperkamp. 2008. "Persistent stress 'deafness': The case of French learners of Spanish". *Cognition* 106, 682–706.

Guion, S.S. 2005. "Knowledge of English word stress patterns in early and late Korean-English bilinguals". *Studies in Second Language Acquisition* 27, 503–533.

Heuven, V.J. van. 2001. *Boven de klanken [Beyond the segments]*. Inaugural address, Universiteit Leiden. Amsterdam: Koninklijke Academie voor Wetenschappen.

Kijak, A. 2009. How stressful is L2 stress? A cross-linguistic study of L2 perception and production of metrical systems. PhD diss., University of Utrecht (published by LOT, Utrecht).

Munro, M.J. & T.M. Derwing. 2006. "The functional load principle in ESL pronunciation instruction: An exploratory study". *System* 34, 520–531.

Peperkamp, S. & E. Dupoux. 2002. "A typological study of stress 'deafness'". *Laboratory Phonology* 7 ed. by C. Gussenhoven & N. Warner, 203–240. Berlin: Mouton de Gruyter.

Trofimovic, P. & W. Baker. 2006. "Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech". *Studies in Second Language Acquisition* 28, 1–30.

Appendix 1: List of target words

	target word	regular stress position	real word
1	abaGIE	–	–
2	abonNEE	–	+
3	abriKOOS	+	+
4	ACtie	+	+
5	aGENda	+	+
6	ALdovas	+	–
7	ALfabet	+	+
8	ALmanak	+	+
9	ALtaar	–	+
10	AMbacht	–	+
11	Atrimon	+	–
12	aviKET	–	–
13	baRAKta	+	–
14	baraTOL	–	–
15	bokaDONT	+	–
16	BRUIloft	–	+
17	buRO	–	+
18	CANnabis	+	+
19	chocoLA	–	+
20	DAKtipauk	–	–
21	detoKAUS	+	–
22	dicTEE	–	+
23	diPLOma	+	+
24	eBANdo	+	–

	target word	regular stress position	real word
25	FEStival	+	+
26	FImoraat	–	–
27	fricanDO	–	+
28	FYsicus	+	+
29	GObilan	+	–
30	goRILla	+	+
31	honkaDIEL	+	–
32	HOrizon	+	+
33	HOSpitaal	–	+
34	iDAlo	+	–
35	kabiNET	–	+
36	kalbisTO	–	–
37	kampiOEN	+	+
38	kaNArie	+	+
39	KAno	+	+
40	kaPEzo	+	–
41	kapiTAAL	+	+
42	karaMEL	–	+
43	KAtapult	–	+
44	kaTOEN	+	+
45	kifuLAN	–	–
46	kleptiLA	–	–
47	koloNEL	–	+
48	koPIE	–	+
49	koZIJN	+	+
50	krisTAL	–	+
51	krokoDIL	–	+
52	LArie	+	+
53	laWAAI	+	+
54	LEKkaduut	–	–
55	lediKANT	+	+
56	LIchaam	–	+
57	makarPEE	–	–
58	maniAK	–	+
59	maroDAX	+	–
60	MASsa	+	+
61	matiNEE	–	+

	target word	regular stress position	real word
62	maTROOS	+	+
63	moeRAS	–	+
64	moraDOF	–	–
65	NIcolaas	–	+
66	nostalGIE	–	+
67	Olifant	–	+
68	OLminek	+	–
69	OORlog	+	+
70	OOievaar	–	+
71	Opa	+	+
72	ovanDIE	–	–
73	palaDROST	+	–
74	PAlaman	+	–
75	paraDIJS	+	+
76	paraPLU	–	+
77	paraSOL	–	+
78	PAridant	–	–
79	paTAT	–	+
80	paWAtie	+	–
81	poluMAS	–	–
82	poRENda	+	–
83	raDIJS	+	+
84	romaTAT	–	–
85	SCHOzafont	–	–
86	seliDAAK	+	–
87	SIEraad	–	+
88	SONtakant	–	–
89	specuLAAS	+	+
90	spiNAzie	+	+
91	stiraGO	–	–
92	toNEEL	+	+
93	Uniform	–	+
94	vaKANtie	+	+
95	VESpodam	+	–
96	VIJand	–	+

Appendix 2: Absolute (and relative) frequency of self-reported proficiency levels in Dutch per L1 subject group, on a scale from 1 (beginner) to 5 (very advanced)

	self-reported DSL proficiency level				
L1	**1**	**2**	**3**	**4**	**5**
French	–	5 (28%)	12 (67%)	1 (5%)	–
Chinese	–	4 (36%)	6 (55%)	1 (9%)	–
Polish	–	1 (11%)	3 (33%)	5 (56%)	–
Finno-Ugric	–	–	3 (60%)	2 (40%)	–
diverse	–	4 (50%)	3 (37%)	1 (13%)	–
total	–	14 (27%)	27 (53%)	10 (20%)	-

Author's address

Johanneke Caspers
Leiden University, LUCL
P. N. van Eyckhof 3
2311 BV Leiden
The Netherlands

J.Caspers@hum.leidenuniv.nl