The production of Dutch word stress by Francophone learners

Marie-Catherine Michaux[1], Johanneke Caspers[2]

marie-catherine.michaux@uclouvain.be, j.caspers@hum.leidenuniv.nl
[1]Université catholique de Louvain – F.R.S.-FNRS
[2]Leiden University Centre for Linguistics

Abstract

This study aims at exploring the production of Dutch word stress by Francophone learners of (Belgian) Dutch. Following other studies, it was hypothesized that participants would show a preference for stressing the final syllable. This hypothesis was confirmed, but the large variability in the data and the lack of agreement between labellers suggest that there is more to it.

1. Introduction

Dutch is a variable-stress language where stress is a lexical property of words (Rietveld & van Heuven 2009) that can be used contrastively (e.g., *voorkomen,* 'to happen', vs. *voorkomen,* 'to prevent'). French does not have contrastive stress: the standard final prominence in isolated words disappears when they are located in non-final position in a larger word group, leaving a word-group final accent (Lacheret-Dujour & Beaugendre 1999; Di Cristo 2000; Rasier 2006). Rather than being contrastive, this 'primary' accent has a demarcative function.

In Dutch, word stress is used as an important cue for word recognition (Cutler 2012; Van Leyden & van Heuven 1996) and Dutch speakers have been shown to be sensitive to mis-stressing (Cutler 2012).

Because French speakers do not have a linguistically encoded prominence at the word level they have sometimes been claimed to be 'stress-deaf' (e.g., Peperkamp & Dupoux 2002; Altmann 2006). However, recent research seems to show that with training Francophones might be able to perceive stress contrasts (Schwab & Llisterri 2012).[1]

In French-speaking Belgium, Dutch is taught as a foreign language in most primary and secondary schools. According to the surveyed students and teachers, pronunciation and prosody, however, are often neglected in Dutch as a Foreign Language (DFL) courses, so that most learners may not be familiar with Dutch word stress.

The production of Dutch word stress has been addressed in small-scaled studies with Francophone learners of Dutch as a second language (DSL) by Caspers and van Santen (2006) and as a foreign language (DFL) by Heiderscheidt and Hiligsmann (2000) and Michaux et al. (2012). Based on the results of these studies it seems clear that the DFL population has to be analysed separately from the DSL one, as the latter group, probably as a result of receiving another type of input (viz. native spoken Dutch), has been found to be more proficient in producing correctly located stress. As for the DFL group it was concluded that learners tend to stick to their L1 pattern, but can also evolve to a penultimate stress (yet not always being the required stress position in Dutch) as time goes by.

This paper investigates DFL word stress production. Following Caspers & van Santen (2006), Michaux et al. (2012) and Schwab & Llisterri (2012), we hypothesized

---

[1] To our knowledge, no study has dealt with the perception of Dutch word stress by Francophone DFL learners, which will be the following step in the current research project.

that Francophone speakers will transfer their L1 final prominence pattern to Dutch words. As familiarity with a phrase-final accent in the L1 might lead to a bias towards final prominence in that position, we also investigated this hypothesis. As a result, non-phrase final words might bear final stress less often in DFL production than phrase-final ones.

2. Method

2.1. Participants

20 DFL learners (age range 19-23, mean age 21.1, 14 females) and 10 native speakers of (Belgian) Dutch (age range 20-51, mean age 28.6, 5 females) took part in the experiment. French was the only mother tongue of the selected DFL speakers.

2.2. Materials

30 existing Dutch three-syllable words were used. They were classified according to the stress rules for simplex words by Trommelen & Zonneveld (1989). The words were split into three canonical stress positions: initial (*pagina,* 'page'), medial (*collega,* 'colleague') and final (*anoniem,* 'anonymous'). Each word was presented thrice in a carrier sentence (*X heb ik gezegd* 'X I said', *Ik heb X gezegd* 'I X said' and *Ik heb gezegd X* 'I said X'), leading to a 90-sentence reading task.

2.3. Procedure

Speakers were recorded individually in a quiet room. Before the recording they filled in a form containing questions about their learner profile (length of Dutch learning, age at start of learning), age and other known/spoken languages.

The trial phase started after an instruction and training session similar to the trial. A Tascam-07 MKII recorder and a Sennheiser PC131 head-set microphone were used.

2.4. Analysis

The data were perceptually labelled independently by a highly-proficient Dutch speaking native French speaker (Labeller 1, Lab 1) and two native Dutch speakers (Lab 2, Lab 3), all of whom were phonetically trained. After listening as often as required to the stimuli, the labellers indicated which syllable they perceived as stressed (1-2-3). Cases of doubt could be expressed as "1?3?", etc.

3. Results

3.1. Interrater agreement

Interrater reliability on the labelled DFL data yielded κ = 0.772 (Lab 1 and Lab 2), κ = 0.674 (Lab 1 and Lab 3) and κ = 0.684 (Lab 2 and Lab 3), all agreements thus being 'substantial' according to Landis & Koch (1977), but not perfect. The agreement between Lab 1 and 2 is slightly higher than the other combinations. The highest κ-scores per participant are found for speaker DFL23 (κ = 0.987, 0.983 and 0.987 resp.), whereas the lowest are found for DFL22 (κ = 0.128, 0.132 and 0.405). Low kappas indicate that the labellers generally did not hear and mark the same syllable as prominent or that they hesitated between several syllables.

For the native speaker group interrater agreement is almost perfect: κ > 0.98 for all pairs of raters.

		No consensus		Syll 1		Syll 2		Syll 3	
Canonical SP	1	27.0	(162)	15.8	(95)	16.5	(99)	40.7	(244)
	2	26.3	(158)	7.2	(43)	30.2	(181)	36.3	(218)
	3	17.4	(104)	9.5	(57)	8.0	(48)	65.1	(390)
		23.6	(424)	10.8	(195)	18.2	(328)	47.4	(852)

Table 1: Percentages (and counts) consensus between labellers for 1st, 2nd and 3rd canonical stress position broken down by perceived stress position.

3.2. Consensus

Based on the labels per labeller, a consensus variable was computed, consensus being reached when per word all labellers marked the same syllable as prominent.

Table 1 shows the consensus values per canonical stress position (SP) for the DFL speakers. The shaded cells contain the cases where canonical and perceived stress concur, leading to correct results. The overall percentage of correct stress amounts to 37.0% (vs. 97.9% for the native group), see 3.3. for more details. According to the hypothesis, canonical SP3 should have been least problematic, which has been born out: SP3 yields the best results (65.1% correct), followed by SP2 (30.2%), which is at chance level, and SP1 (15.8%). On the whole there is a preference for syllable 3 (47.4%) regardless of the canonical SP, but there is also substantial variation in the remainder of the data. While there is a clearer preference for syllable 3 in the other cases, SP2 shows a different pattern: both syllable 2 and 3 yield approx. 30% of consensus (chance).

In roughly 25% of the cases with SP1 and SP2, no agreement (i.e. 'no consensus' in the table) was reached between the labellers, while this percentage reaches 17.4% for SP3. Strikingly, 'no consensus' very often has the second highest frequency after consensus on the 3rd syllable. This either means that the labellers could not determine which syllable was stressed or that they perceived different or several syllables as stressed.

3.3. Percentage of correct stress

The percentage of correct stress in the DFL material being 37.0% (see 3.2.), it seems safe to claim that the DFL speakers in this study had not mastered Dutch word stress yet.
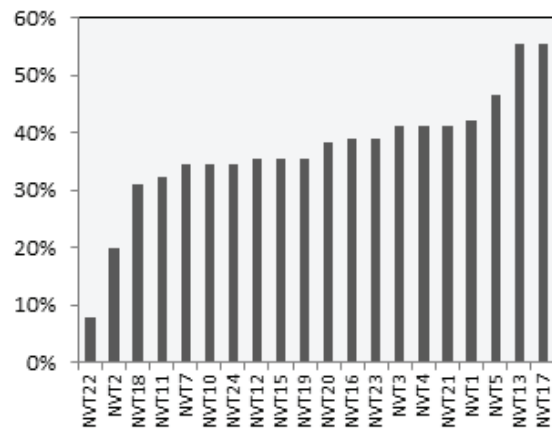

Figure 1: Percentages of correct stress per speaker, ranked from lowest to highest.

Figure 1 shows the percentage of correct word stress per speaker. The highest scores are reached by DFL13 and DFL17, but do not exceed 60.0%. The lowest score is found for DFL22 who unsurprisingly is also the participant with the lowest κ-values. However, the participant with the highest κ-scores (DFL23) does not achieve the highest correctness score (38.9%), meaning that a clear stress realisation does not necessarily imply a correct stress location. The majority of the participants reach between 30% and 40% correctness, which again shows how little grasp the speakers had of Dutch word stress.

A repeated measures ANOVA (with Greenhouse Geisser correction) of the percentage correct, aggregated over stimulus words, with canonical SP and word position in the sentence as within-subjects factors and L1 as between-subjects factor, shows an effect of L1, $F_{(1,28)} = 324.1$ (p<.001) and canonical SP, $F_{(1.6,45.3)} = 8.5$ (p<.001), and an interaction between canonical SP and L1 $(F_{(1.6,45.3)} = 8.3$ (p<.001). This means that the DFL production varies a lot more for different canonical SPs than the native production does. Pairwise comparisons reveal that the effect of SP is caused by the difference between SP3 and the other SPs. Contrary to our hypothesis, there is no effect of the position in the sentence $(F_{(2.0,55.0)} =$

2.2, ins.), meaning that there is a bias towards syllable 3 at the word level (see 3.2.), but not at the phrase level. There is no interaction between L1 and position in sentence ($F_{(2.0,55.0)}= 1.6$, ins.), and between position in sentence, L1 and SP ($F_{(3.4,95.4)}< 1$, ins.).
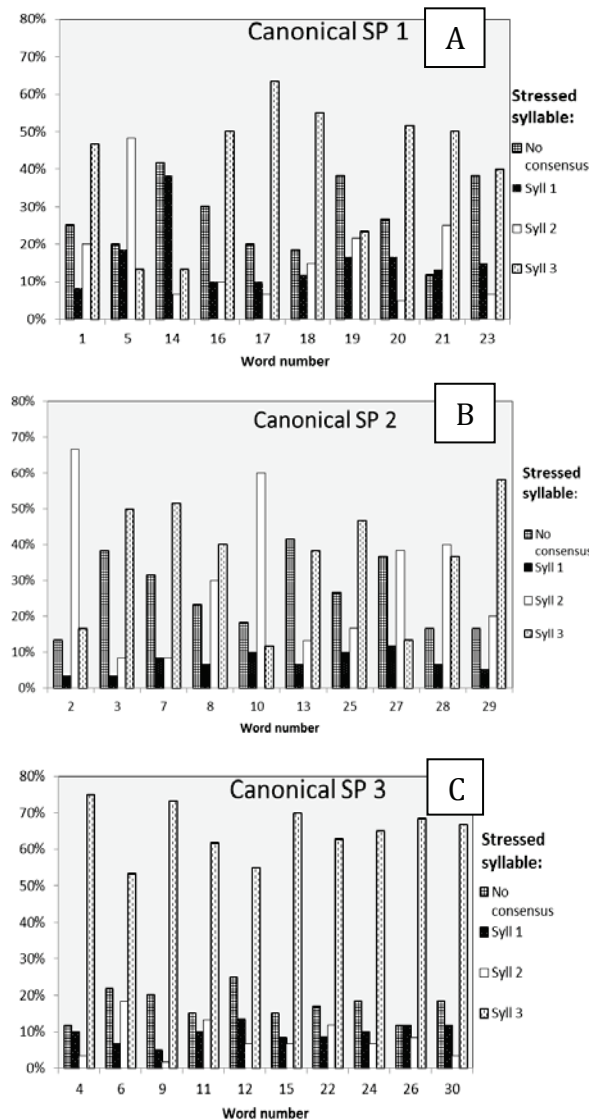

Figure 2: Stressed syllable per word and canonical SP1 (A), SP2 (B) and SP3 (C).

3.4. Variability in the data

Figure 2 shows the stressed syllable (based on consensus) per canonical SP and stimulus word. SP3 (panel C) shows less variability ($\chi^2_{(27)}= 31.24$, ins.) than categories SP1

(panel A) and SP2 (panel B) where the stressed syllable seems to vary more depending on the stimulus word pronounced (resp. $\chi^2_{(27)}= 144.8$, p<.001 and $\chi^2_{(27)}= 147.5$, p<.001).

Given the variation in the data, the words were further analysed according to their form and stress pattern similarity to French. The words were split into four categories: (1) same form, same prominence location, e.g., *chocola*, Fr. 'chocolat' ('chocolate'), (2) same form, different prominence location, e.g., *marathon*, (3) different form, same prominence location, e.g., *abrikoos*, Fr. 'abricot' ('apricot'), (4) different form, different prominence location, e.g., *augustus*, Fr. 'août' ('August'). Figure 3 shows the distribution of stressed syllables per category.
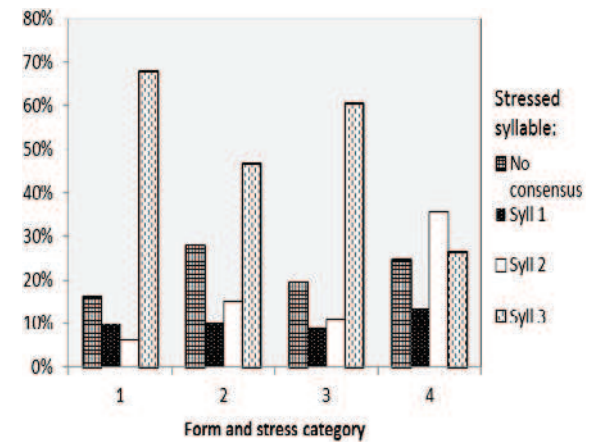

Figure 3: Percentage of stressed syllable per form and stress category (see text).

Categories (2) and (4) yield proportionally more 'no consensus' (resp. 28.1%, 24.6%) than other categories. Furthermore, category (4) yields stress more often on syllable 2 (35.6%) than other categories ((1) 6.7%, (2) 15.1%, (3) 10.0%). In contrast, stress on syllable 3 occurs much less often in category (4) (26.3%) than in other cases ((1) 67.8%, (2) 46.7%, (3) 60.5%). Category (4) comprises words such as *collega* (Fr. 'collègue', En. 'colleague') and *augustus*. One could argue that a correct stress location on the second syllable is related to

the high frequency of those words. If anything, this result also suggests that words formally less similar to French should be viewed separately, as they seem to lead to another stressing behaviour and might be less prone to the transfer of the French final pattern.

4. Conclusion and discussion

The low scores of the participants clearly show poor knowledge of Dutch word stress. On the whole, the speakers relied on their final L1 pattern to stress the stimuli regardless of the L2 canonical stress position. This globally supports our hypothesis.

However, all cases taken together, the highest percentage of final stress amounted to 47.7%, which means that in the remainder of the cases there is either variability in the realized stress position or lack of agreement between the labellers. The second most frequent labelling after syllable 3 was not any other stress position, but the category 'no consensus'; after the labelling process the labellers all stated how unstable they had found the data, some of the stimuli bearing several prominences, sometimes felt to be rendered with different cues. In the aftermath of this production study all production data will be acoustically analysed to find the reason for this perceptual uncertainty. Hypothetically, if several syllables were indeed stressed, one could argue that some kind of 'arc accentuel' has been produced (Di Cristo 1998) or that the participants showed an unsteady behaviour because of the lack of a mentally stored stress. However, as some words (category (4)) showed a preference towards stress on the second syllable, one could also argue that in some cases the participants realised that there is a prosodic difference between Dutch and French. This would be in agreement with the findings of Michaux et al. (2012).

Finally, contrary to our expectations, a final location of the word within the sentence did not yield final stress more often, which forces us to reject our second hypothesis.

Acknowledgments

References

Altmann, H. (2006). *The Perception and Production of Second Language Stress: A Crosslinguistic Experimental Study*. Diss., University of Delaware.

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?[Dutch by French and Chinese learners. L1 influence on stress position in Dutch as a second language] *Nederlandse Taalkunde, 11*(4), pp. 289-318.

Cutler, A. (2012). *Native Listening. Language Experience and the Recognition of Spoken Words*. MA: MIT Press, Cambridge.

Di Cristo, A. (1998). Intonation in French. D. Hirst & A. Di Cristo (eds.), *Intonation Systems: A Survey of Twenty Languages,* CUP, pp. 195-218.

Di Cristo, A. (2000). Vers une modélisation de l'accentuation du français (seconde partie). *French Language Studies* 10, pp. 27-44.

Heiderscheidt, S. & P. Hiligsmann (2000). De accentuering in de tussentaal van Franstalige leerders van het Nederlands. *Leuvense Bijdragen,* 89(1/2), pp.17-131.

Lacheret-Dujour, A. & F. Beaugendre (1999). *La prosodie du français*. CNRS Editions, Paris.

Landis, J.R. & G.G. Koch (1977). The measurement of observer agreement for categorical data. *Biometrics* 33(1), pp. 159-174.

Leyden, K., van & V.J. van Heuven (1996). Lexical stress and spoken word recognition: Dutch versus English. *Linguistics in the Netherlands 1996,* pp.159-170.

Michaux, M.-C., Ph. Hiligsmann & L. Rasier (2012). Het klemtoonpatroon in de tussentaal van Franstalige leerders van het Nederlands [The stress pattern in the interlanguage of francophone learners of Dutch]. *XII. Internationaler Germanistenkongress* 2010, Warsaw, pp. 321-332.

Peperkamp, S. & E. Dupoux (2002). A typologocial stress 'deafness'. In C. Gussenhoven & N.

Warner, (Ed.), *Laboratory Phonology 7*, Mouton de Gruyter, Berlin, pp. 203-240.

Rasier, L. (2006), *Prosodie en vreemdetaal-verwerving. Accentdistributie in het Frans en Nederlands als vreemde taal.* Diss. UCLouvain.

Rietveld, A. C. M. & V.J. van Heuven (2009). *Algemene Fonetiek [General Phonetics].* Bussum, Coutinho.

Schwab, S. & J. Llisterri (2012). Are French speakers able to learn to perceive lexical stress contrasts? *17th ICPhS,* City University of Hong Kong, pp. 1774-1777.

Trommelen, M. & W. Zonneveld (1989). *Klemtoon en metrische fonologie [Stress and metrical phonology].* Muiderberg, Coutinho.