

Lexical Stress, Phrasal Accent and Prosodic Boundary in the Realization of Domain-initial Stops in Dutch

Taehong Cho

Max Planck Institute for Psycholinguistics
Postbus 310, 6500 AH, Nijmegen, The Netherlands
E-mail: Taehong.Cho@mpi.nl

ABSTRACT

This study examines the effects of prosodic boundaries, lexical stress, and phrasal accent on the acoustic realization of stops (/t, d/) in Dutch, with special attention paid to language-specificity in the phonetics-prosody interface. The results obtained from various acoustic measures show systematic phonetic variations in the production of /t d/ as a function of prosodic position, which may be interpreted as being due to prosodically-conditioned articulatory strengthening. Shorter VOTs were found for the voiceless stop /t/ in prosodically stronger locations (as opposed to longer VOTs in this position in English). The results suggest that prosodically-driven phonetic realization is bounded by a language-specific phonological feature system.

1. INTRODUCTION

A large body of studies on the phonetics-prosody interface has provided converging evidence that an utterance is produced in an organizational frame of hierarchically-nested prosodic structure, which reflects the grouping and relative salience of phonological units. More recently, quite a few studies have suggested that prosodic structure can be phonetically marked at least in part by consonantal strengthening associated with prosodically strong locations [1,2,3]. For example, voiceless stops are generally produced with longer VOTs at onsets of larger prosodic constituents than at onsets of smaller ones [2,4]. However, few studies have considered multiple prosodic factors concurrently in examining prosodically-conditioned consonantal realization.

The present study investigates the effects of three prosodic factors (i.e., lexical stress, phrasal accent and prosodic boundary) on the acoustic realization of stops (/t d/) in Dutch in order to gain better insight into two inter-related aspects of the phonetics-prosody interface: (1) how the phonetic realization of stops is conditioned by prosodic structure and (2) from a different perspective, how high-level prosodic structure itself is signaled by systematicity found in phonetic details associated with consonantal production. Specifically, this study will consider the language-specificity of the phonetics-prosody interface by examining how voiceless stops in Dutch are acoustically realized in comparison with English stops. Given that English voiceless stops are generally produced with longer VOTs in prosodically stronger positions (i.e., lexically-

stressed, phrase-accented and domain-initial syllables), this study tests whether Dutch voiceless stops are produced in much the same way.

2. METHODS

Eleven Dutch speakers read various sentences (Table 1) designed to induce different boundary types as well as different phrasal accents .

-
- a. Utterance:** Roel ging naar MIJN opa, **T**acitus DE VRIES was gisteren jarig.
(*Roel went to my grandpa. It was Tacitus de Vries' birthday.*)
- b. Vocative:** ZEG eens opa, **T**acitus DE VRIES is net aangekomen.
(*Say, Grandpa, Tacitus de Vries has just arrived.*)
- c. List:** Vandaag gingen MIJN opa, **T**acitus DE VRIES en Zara de trap af.
(*Today, my grandpa, Tacitus de Vries and Zara went down the stairs.*)
- d. Verb-Object Inversion:** Dan gaat MIJN opa **T**acitus DE VRIES helpen inpakken.
(*Then my grandpa is going to help Tacitus de Vries to pack.*)
- e. Word:** VANMORGEN ging opa **T**acitus DE VRIES de trap af.
(*This morning, Grandpa Tacitus de Vries went down the stairs*)
-

Table 1. An example set of sentences containing /t/.

The target consonant was varied to be /t d/ in either stressed or unstressed syllable as in *T*acitus, *T*amára, *D*aniel and *D*anıla. For accent-placements, two different versions of each sentence were produced: one with accent falling on the target word (e.g. *Tacitus*, the first name) and one with accent on the next word (e.g. *de Vries*, the last name). In order to induce the desired variety of accent-placement patterns, words to be accented were written in bold upper case in the script, and speakers were introduced to discourse situations in which the target sentences may occur. Finally, five different sentence types (Table 1a-e) were used so that the target consonants occurred at beginnings of prosodic constituents of different size.

Before the actual data collection, speakers practiced saying the speech materials for 20-30 minutes to produce the intended renditions as naturally as possible. Each sentence was repeated three times in a pseudo-randomized order. In total, 1320 sentence tokens were collected (2 consonants x 2 stress levels x 2 accent levels x 5 sentence types x 11 speakers x 3 repetitions). The collected sentences were divided into three prosodic groups

according to their prosodic boundary strengths (BSs):

- BS3: boundary tone and pause (strongest)
- BS2: boundary tone and no pause (intermediate)
- BS1: no boundary tone and no pause (weakest)

BS was determined following de Pijper & Sanderma [5] who showed that perceived boundary strength is stronger when there is a boundary tone together with a pause (BS3) and weaker when there is neither a boundary tone nor a pause (BS1). Since the acoustic signal does not differentiate a pause and a stop closure, the pause was defined separately for each speaker as any period of silence exceeding a double of the mean closure duration of the stop produced in the Word sentences (see Table 1e).

BS1 and BS3 are roughly equivalent to the Prosodic Word boundary and the Intonational Phrase (IP) boundary, respectively in the theory of prosodic organization [6]. Sentences belonging to the BS2 category (with no pause), however, could have been further divided into IP and a smaller phrase (e.g., the Intermediate Phrase or the Phonological Phrase). However, because of the lack of a widely agreed-upon prosodic transcription system in Dutch, the present study follows the Boundary Strength model [5] without further differentiating between IP and a smaller phrase within the BS2 category.

Acoustic measurements included (a) duration of preboundary syllables (e.g., /pa#/ in *opa*, where # = a prosodic boundary), (b) closure duration of /t d/ (e.g., /#ta/ in *Tacitus*) (excluding /t d/ in BS3 because the pause and the stop closure cannot be differentiated), (c) voicing duration during /d/, (d) VOT for /t/, (e) postboundary vowel duration in #CV (e.g., /a/ in *Tacitus*), (f) burst energy (dB) for initial stops, and (g) the center of gravity at the release burst. The last two measurements were obtained from FFT spectra using a 256-point (12.8 ms) hamming window centered at the stop release.

3. RESULTS

3.1 Preboundary Lengthening

As preboundary lengthening has widely been considered as one of the primary phonetic correlates of prosodic structure, let us first examine whether the three-way prosodic grouping based on BS can be further corroborated by preboundary lengthening. As shown in Figure 1, results of repeated measures ANOVAs conducted separately for each consonant show that there is a reliable three-way distinction in preboundary lengthening that corresponds to the three-way prosodic grouping ($p < 0.05$, Bonferroni/Dunn).

3.2 /t/

The results for /t/ are summarized in Table 2.

Closure Duration ([t]_DUR). There are main effects of all three factors (Stress, Accent and BS) on /t]_DUR. It is significantly longer in stressed and accented syllables and at a stronger boundary (e.g., BS2) than in unstressed and unaccented syllables and at a weaker boundary (e.g., BS1), respectively.

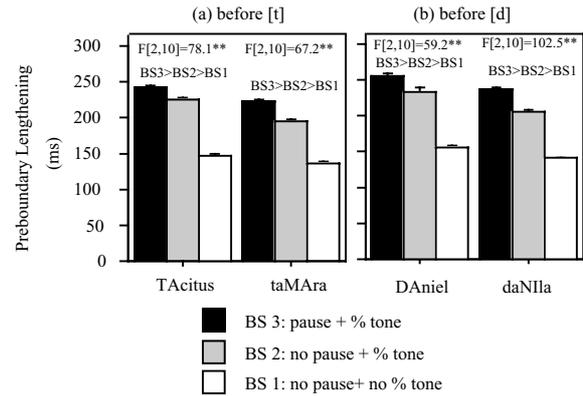


Figure 1. Preboundary lengthening of CV#. ** = $p < 0.001$; > = $p < 0.05$.

	Stress	Accent	BS
[t]_Dur (ms)	STR > UNS (99.4, 69.7 ms) F(1,10)=90.8**	ACC > UNA (93.9, 72.2 ms) F(1,10)=27.8**	BS2 > BS1 (100.1, 69.6 ms) F(1,10)=167.2**
No interactions			
VOT (ms)	STR < UNS (23.1, 37.8 ms) F(1,10)=60.1**	ACC < UNA (28.5, 32.3 ms) F(1,10)=20.2**	BS3 < {BS2=BS1} (27.6, 31.5, 32.2 ms) F(2, 20)=8.1**
STR x BS: F(1.7,16.5)=6.773** (see Figure 2)			
Burst (dB)	n.s.	n.s.	n.s.
No interactions			
COG (Hz)	n.s.	n.s.	BS3 > BS1 (5194, 5259, 5292) F(2,20)=10.74**
No interactions			
V Dur (ms)	STR > UNS F=477.2**	ACC > UNA F=100.7**	n.s.
STR x BS: F(1.8,17.7)=7.67**			

Table 2. ANOVAs for /t/. * = $p < 0.05$ and ** = $p < 0.01$.

VOT. There are main effects of all three factors on VOT, showing *shorter* VOTs for /t/s in stressed and accented syllables and at a stronger boundary (BS3). As shown in Figure 2a, VOT is shortest when accented *and* stressed and it is longest when unaccented *and* unstressed, showing a cumulative effect of Stress and Accent on VOT. There is a significant Stress x BS interaction as shown in Figure 2b. The effect of BS on VOT is more robust when /t/ is unstressed than stressed: Posthoc tests show that when stressed, VOT is significantly shorter for BS3 than for BS2 with no reliable difference between BS3 and BS1, whereas when unstressed, it is significantly shorter for BS3 than for both BS2 and BS1.

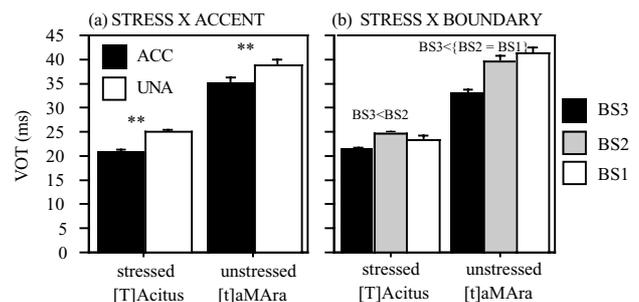


Figure 2. Variations in VOT. ** = $p < 0.001$; > = $p < 0.05$.

Burst Energy (dB). There is neither a main effect nor any interactions among the factors.

Center of Gravity (COG, Hz). Neither Stress nor Accent influences COG, but there is an effect of Boundary Strength: COG is significantly higher for BS3 than for BS1.

Vowel Duration in /ta/. There are main effects of Stress and Accent on Vowel Duration: /a/ in /ta/ is longer when it is either stressed or accented, with no Stress x Accent interaction. Boundary Strength shows no main effect on Vowel Duration, but there is a Stress x Boundary Strength interaction due to the fact that Boundary Strength influences vowel duration, showing a pattern of BS3>BS1 only when the syllable is unstressed.

3.3 /d/

The results for /d/ are summarized in Table 3.

Closure Duration ([d]_DUR). There are main effects of all three factors on [d]_DUR, showing longer closure duration when stressed (vs. unstressed), accented (vs. unaccented) and at a stronger prosodic boundary (B2>B1). However, there is a Stress x Accent interaction: When stressed, accented /d/ is longer than unaccented /d/, but not significantly so when /d/ is unstressed. This suggests that unstressed /d/ is less sensitive to accentuation, which is presumably because phrasal accent generally influences the articulation of stressed syllables.

	Stress	Accent	BS
[d]_Dur (ms)	STR > UNS F(1,10)=12.5**	ACC > UNA F(1,10)=32.6**	BS2>BS1 F(1,10)=51.3**
	STR x ACC: F(1,10)=35.1**		
Voicing Dur	STR > UNS F(1,10)=7.2*	ACC > UNA F(1,10)=15.6**	BS3<{BS2=BS1} F(2,20)=37.3**
	STR x ACC: F(1,10)=58.6**		
Burst (dB)	<i>n.s.</i>	<i>n.s.</i>	BS2 > BS1 F(1.7,17)=4.6*
	STR x ACC: F(1,10)=5.5*		
COG (Hz)	<i>n.s.</i>	<i>n.s.</i>	BS2 > BS1 F(1.6,15.9)=4.7*
	No interactions		
V Dur (ms)	STR>UNS F(1,10)=174.8**	ACC>UNA F(1,10)=99.6**	<i>n.s.</i>
	STR x ACC: F(1,10)=43.3**		

Table 3. ANOVAs for /d/. * = p<0.05; ** = p<0.01.

Voicing Duration. There are main effects of all three factors on Voicing Duration, but they differ in terms of directionality. Voicing Duration is longer when stressed (vs. unstressed) and when accented (vs. unaccented), but *shorter* at a stronger prosodic boundary (BS3>BS2=BS1). (See the next section for the discussion of the shorter voicing duration at a stronger boundary.) There is also a Stress x Accent interaction that comes from the fact that the stressed/unstressed distinction in voicing duration is maintained only when /d/ is accented.

Burst Energy (dB). There are no main effects of Stress and Accent on Burst Energy. A significant Stress x Accent interaction, however, shows that stressed /d/ has greater Burst Energy than unstressed /d/ mainly when accented. Again, the stressed/unstressed distinction in Burst Energy

appears to be reinforced by accentuation. As regards BS, there is a main effect showing larger Burst Energy for a stronger prosodic boundary (B2>B1).

Center of Gravity (COG, Hz). As in Burst Energy, while no main effects are observed for the Stress and Accent factors, there is a main effect of Boundary Strength on COG showing a higher centroid frequency at the release for a stronger prosodic boundary (BS2>BS1).

Vowel Duration in /da/. /a/ in /da/ is significantly longer when stressed (vs. unstressed) and accented (vs. unaccented). However, there is a Stress x Accent interaction which is again due to a more robust stress-induced difference under accent. As for BS, as was found for /ta/, there is no BS effect on Vowel Duration.

4. SUMMARY AND DISCUSSION

a. Both /t/ and /d/ are produced with longer closure duration not only in stressed and accented syllables, but also at a stronger prosodic boundary (when considering only BS2 and BS1 because of the involvement of the pause for BS3). These results are compatible with prosodically-conditioned strengthening phenomena that have been reported in other languages [1,2,3,7].

b. The effect of prosodically-conditioned strengthening is further supported by greater burst energy at a stronger prosodic boundary. There were, however, no stress or accent effects on Burst Energy. One exception to this was that a significant stress/unstressed distinction in Burst Energy (greater burst energy when stressed) was observed for /d/ when the target word received phrasal accent, suggesting an augmentation of the stress effect under accentuation.

c. The Center of Gravity measured at the stop release does not reliably mark stressed and accented syllables. However, both /t/ and /d/ show significantly higher COG at a stronger prosodic boundary. The higher COG at the stop release likely indicates decreased length of the front cavity which may be ascribable either to larger linguopalatal contact during stop closure or to tongue fronting or to both. Given that the longer closure duration of coronal consonants at a stronger prosodic boundary is generally associated with larger linguopalatal contact [2], the heightened COG at a stronger prosodic boundary can be interpreted as being attributable to articulatory strengthening as would be reflected in larger linguopalatal contact.

d. Vowel duration is clearly lengthened in both stressed and accented syllables, but not robustly so at a prosodically stronger boundary. However, there is a Stress x BS interaction for /ta/, such that vowel duration is longer at a stronger prosodic boundary when /ta/ is *unstressed*. This has some implications for the assumption that the boundary-induced strengthening effect (aka domain-initial strengthening) is localized to beginnings of prosodic domains, i.e., to C in #CV. It may be the case that V in #CV is further away from the boundary, such that the

strengthening effect wanes into the vowel. But, the results found in this study suggest that domain-initial strengthening may include the lengthening effect on V in #CV, especially when the syllable is unstressed. One possible explanation is that vowels are already lengthened when stressed, leaving less room for temporal expansion. But when vowels are unstressed, there is no such a ceiling effect, so the (boundary-induced) temporal expansion of vowels may be possible.

e. For /d/, voicing duration is longer when accented or stressed, but *shorter* at a higher prosodic boundary. The shortened voicing duration may be interpreted as an enhancement of syntagmatic (or CV) contrast at a higher prosodic boundary: Decreased voicing at a higher boundary would make a voiced consonant more consonant-like (less sonorous), enhancing a CV contrast.

f. For /d/, the fact that the stressed/unstressed distinction in voicing duration is maintained only when /d/ is accented suggests that the stressed/unstressed distinction is augmented by accentuation. Voicing requires an active laryngeal adduction gesture along with an adequate aerodynamic condition, and accentuation is likely to bring about an increased articulatory/aerodynamic force to facilitate the voicing gesture.

g. The voiceless stop /t/ is produced with *shorter* VOT in all prosodically strong locations. This pattern is the opposite of the VOT pattern found in English [4,8]. However, the opposite patterns can still be accounted for by the same principle of prosodically-conditioned articulatory strengthening if we take into account the effect of the language-specific phonological feature system on the phonetic realization of segments. Both Dutch and English have a two-way phonological contrast in stops (voiceless vs. voiced). In order to maintain such a phonemic contrast, we may only need the feature [+/-voiced] for both languages. Nevertheless, Dutch voiceless stops are generally produced with shorter VOTs (less aspiration) than English ones. As Keating [9] proposed, in order to account for this kind of cross-linguistic variation, a secondary feature is necessary such as [+/-spread glottis]: English voiceless stops may be specified with [+s.g] (which is phonetically implemented by the glottal abduction (opening) gesture, resulting in relatively long VOTs) whereas Dutch voiceless stops may be specified with [-s.g] (which is phonetically implemented by the glottal adduction (closing) gesture, resulting in relatively short VOTs). Given such cross-linguistic differences both at the phonological and phonetic implementation levels, it follows that prosodically-conditioned articulatory strengthening in English is reflected in the strengthening of the glottal abduction gesture, resulting in longer VOTs, whereas in Dutch it is reflected in the strengthening of the glottal adduction gesture, resulting in shorter VOTs. Furthermore, the elongated VOT in English can be interpreted as an enhancement of [+s.g.] and the shortened VOT in Dutch as an enhancement of [-s.g.].

5. CONCLUSION

The present study has shown that the phonetic realization of /t/ and /d/ varies systematically with the prosodic position in which they occur. Such prosodically-driven phonetic patterns can be seen as signaling hierarchically-nested structure of speech organization, from which listeners may ultimately benefit in speech comprehension [10]. Furthermore, the results illuminate how prosodically-conditioned articulatory strengthening is bounded by language-specific phonological feature system and phonetic implementation. One extremity in phonetic values in the continuum of a phonetic parameter (e.g., VOT) may mean strengthening in one language but weakening in another.

REFERENCES

- [1] C. Fougeron & P.A. Keating. "Articulatory strengthening at edges of prosodic domains." *Journal of the Acoustical Society of America*, vol. 101 (6), pp. 3728-3740, 1997.
- [2] T. Cho & P.A. Keating. "Articulatory and acoustic studies of domain-initial strengthening in Korean." *Journal of Phonetics*, vol. 29, pp. 155-190, 2001.
- [3] P.A. Keating, T. Cho, C. Fougeron & C. Hsu. "Domain-initial strengthening in four languages," in J. Local, R. Ogden, and R. Temple, *Papers in Laboratory Phonology VI*, Cambridge, UK: Cambridge University Press, In press.
- [4] J. Pierrehumbert & D. Talkin. "Lenition of /h/ and glottal stop," in G. Docherty and D.R. Ladd, *Papers in Laboratory Phonology II*, Cambridge, UK: Cambridge University Press, pp. 90-117, 1992.
- [5] J.R. de Pijper and A.A. Sanderman. "On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues." *Journal of the Acoustical Society of America*, vol. 96 (4), pp. 2037-2047, 1994.
- [6] M.E. Beckman & J. Pierrehumbert. "Intonational structure in Japanese and English." *Phonology Yearbook*, vol. 3, pp. 255-309, 1986.
- [7] T. Cho. *The effects of prosody on Articulation in English*. New York, NY: Routledge, 2002.
- [8] L. Lisker and A. Abramson. "Some effects of context on voice onset time in English stops." *Language and Speech*, vol. 10, pp. 1-28, 1967.
- [9] P. Keating. "Phonetic and Phonological Representation of stop consonant voicing." *Language*, vol. 60, pp. 286-319, 1984.
- [10] J. McQueen and T. Cho. "The use of domain-initial strengthening in segmentation of continuous English speech." *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 2003.