Voicing of word-final obstruent in English connected speech

Dźwięczność wygłosowych obstruentów w mowie ciągłej w języku angielskim

Wiktor Gonet

Maria Curie-Skłodowska University, Lublin
gonet@umcs.pl

ABSTRACT

Implementation of the voicing of English word-final obstruents constitutes a serious problem for foreign speakers of English whose native languages organize the phonological 'voicing' contrast differently. Polish teachers and learners of English implement too few and too sharply-cut rules whose application produces unnatural rendering of the English voicing distinction, with overemphasis on the retention of voicing in word-final obstruents. A review of literature based mostly on the study of isolated words has revealed a lacuna concerning the description of the behaviour of English word-final lenis obstruents in connected speech. It is to this issue that the present paper is addressed.

STRESZCZENIE

Właściwa realizacja dźwięczności angielskich wygłosowych obstruentów (tj. spółgłosek niesonornych) stanowi poważny problem dla osób, w których języku ojczystym kontrast dźwięczności jest zorganizowany inaczej niż w angielskim. W procesie glottodydaktycznym dźwięczność omawianych głosek opisywana jest za pomocą niewystarczającej liczby zbyt ostro sformułowanych zasad, z naciskiem na zachowanie dźwięczności tych głosek w wygłosie. Przegląd opisów angielskiej fonetyki wykazuje ponadto, iż opierają sie one głównie na badaniach pojedynczych słów, bez uwzględnienia kontekstu. Niniejsza praca stawia sobie za cel uzupełnienie opisu realizacji dźwięczności wygłosowych angielskich obstruentów w zależności od kontekstu.

1. Introduction

The implementation of the voicing of the English obstruents constitutes a serious problem for foreign speakers of English whose native languages organize the phonological 'voicing' contrast differently. In English one has to do with a three-way implementation of phonological voicing, i.e. voiceless, partially voiced and fully voiced, while in Polish the contrast is two-way and fully neutralized in absolute word-final position. As was argued in [1], while in Polish the auditorily more prominent member of the opposition is the voiced sound, in English the more prominent opposition member is the voiceless sound, especially if it is aspirated, so the English voiced sound should rather be defined as unaspirated and weak.

In both languages the voicing of the phonologically voiced obstruents varies depending on the context. In Polish there is a rule of word-final devoicing, but its activity is counterweighed by voicing retention in connected speech whose degree depends on dialectal variability. In the conviction of the present author the actual realization of English word-final lenis obstruents has not so far been described satisfactorily. Polish teachers and learners of English, in a desire to counteract the Polish word-final devoicing, attempt to retain full voicing in English word-final obstruents, irrespective of the following context. By implementing such simplified rules, teaching English pronunciation produces unnatural rendering of the English voicing distinction. A review of the literature that describes mostly isolated words has revealed a lacuna concerning the description of the behaviour of English word-final phonologically voiced obstruents in connected speech. In this pedagogical context, a detailed description of English obstruent voicing in word-final position is necessary to set up a standard on which foreign teachers and learners should rely.

Vocal fold vibration is by no means the only aspect of English pronunciation that is responsible for the phonological voiced-voiceless contrast. Quite to the contrary, besides the actual primary vocal fold vibration, the implementation of the phonological feature of 'voicing' involves a bundle of 'secondary' phonetic features (such as lack of aspiration, relatively weak force of articulation, increased durations of preceding vowels and sonorants) that, unlike voicing, remain stable irrespective of the context. It has been convincingly argued in a number of sources that those 'secondary' properties of the voiced obstruents are in fact more important as perception clues than the actual vocal fold vibration since native speakers of English "don't really pay much attention to the actual voicing in the segment itself" [2]. Not questioning this stand, it still remains to be shown, not only in the sphere of English pronunciation pedagogy, how actually voicing is implemented in English word-final lenes.

2. Goals

The problem of the implementation of voicing specifically concerns the rendering of word-final voiced obstruents in connected speech, in which they can be immediately followed by (i) a pause, (ii) a voiceless sound and (iii) a voiced sound. This question will be approached in the present paper in a quantiative framework, where the informal phrase 'the amount of voicing' refers to the duration of the vocal fold vibration segment extending from the previous voiced sound into the constriction (narrowing or full closure). The goal of the paper is to find answers to the following questions:

1. Is the amount of voicing intrinsically related to the manner of articulation?
2. Does the preceding or following context exert any effect on the duration of the voicing interval?

3. Previous studies

In a number of popular textbooks of English pronunciation the generalizations are specified in imprecise terms. For instance, Davenport and Hannahs state [3:p24] (emphasis added):

> Only in instances like those above [i.e. lobbing, lagging, ladder, saving], i.e. between two other voiced sounds is an English voiced stop fully voiced. Elsewhere, such stops are likely to be wholly or partly devoiced.

Ball and Rahilly [4:pp72–73] (emphasis added) observe that the voicing of lenis obstruents is determined by the position in the word:

> Voiced stops, […] may be subject to a certain amount of loss of voicing (i.e. voicelessness, often termed 'devoicing') during their production. This is common in English, for example, in word-initial and word-final positions. […] Indeed in English word-initial and word-final stops for some speakers may be totally devoiced.

Catford [5:p46] (emphasis added) maintains:

> In English, final 'voiced stops' […] are generally only partly voiced. Thus, the final stops of such words as ebb, Ed, egg, may be almost completely voiceless […].

Roach [6:p35] (emphasis added) asserts that:

> Final b, d, g normally have little voicing; if there is voicing, it is at the beginning of the compression phase […]. The difference between p, t, k and b, d, g is primarily the fact that vowels preceding p, t, k are much shorter.

The rules are formulated in a way that allows considerable variability, though the authors mentioned above do not specify the frequency of voicing retention with regard to intra-speaker and inter-speaker variability, or durations of voiced segments.

While so far the domain of the devoicing process was the word, there are two authors who extend the domain of voicing across word boundaries. Gimson [7] presents a number of exemplary phrases (e.g. 'egg and') that show that in deciding on the amount of voicing in a plosive a broader context is to be taken into account. Gimson [7] also takes into account silence before and after the obstruent:

> In initial and especially in final positions, i.e. following or preceding silence, /b, d, g/, while remaining lenis, may be only partially voiced or completely voiceless, e.g. in bill, done, game, cub, lid, bag. (p. 160, emphasis added).

In this position, according to the same author, the vocal cord vibration finishes in the first portion of the compression stage, or there is no voicing at all in this stage. Gimson [7:p160] (emphasis added) continues to say:

> [E]ven in the intervocalic positions […] /b, d, g/ may sometimes be subject to devoicing, particularly where a word boundary is involved.

Thus a devoicing context is constituted by silence, while the non-devoicing context is that between two voiced sounds, including the situations in which the preceding word ends, and the next word starts, with voiced sounds. This means that the rule 'preserve the voicing of the lenis obstruents between voiced sounds' can be overridden by the rule 'devoice lenis obstruents that are word-initial or word-final'.

The view on the voicing of lenis obstruents presented by Jassem [8:pp203–207] is still broader in terms of contextual possibilities and so far the most adequate: the devoicing of lenis obstruents does not only depend on word-initial or word-final position, with a pause preceding or following, but also on the presence of voiceless sounds

following or preceding lenis obstruents across word boundary. With regard to devoicing, Jassem [8:pp203–207] observes that obstruents have first part voiced, second part voiceless if the occlusive is preceded by a voiced phone and followed by a voiceless phone or a pause. Additionally, they can be voiceless throughout, partly in free variation with both types of the partially voiced plosives.

It can be concluded that, according to most authors, the voicing of obstruents is determined by reference to the position in the word. Gimson [7] also mentions that voicing is preserved between two voiced sounds of which the second belongs to the next word, while devoicing can also happen before silence. Only Jassem [8] refers to the quality of the following word-initial sounds.

In [9], a study of lenis stops (plosives and affricates, in figure captions abbreviated to /d/) and fricatives (abbreviated as /z/) was undertaken, with three qualitative designations (fully voiced, partially voiced, and voiceless) in two following types of context – 'voicing-promoting' (vowels, voiced obstruents) and 'voicing-impeding (pre-pausal, voiceless obstruent). The study revealed that in stops the incidence of full voicing increased from 8% before a pause, through 26% if followed by a voiceless sound, to 83% if followed by a voiced sound (Figure 1).

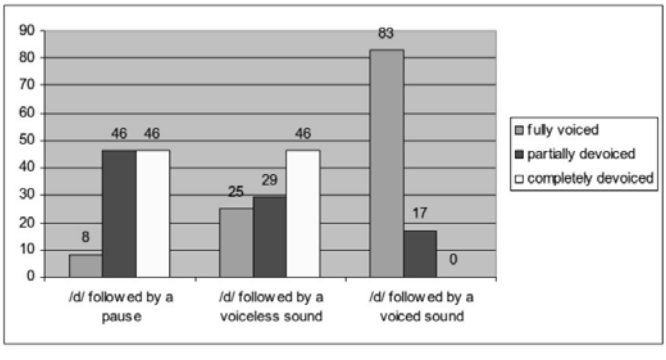

Figure 1: Voicing of word-final plosives depending on the following context.

In fricatives, the increase of voicing was not so well manifested when the figures for the voicing-impeding environment (4%) were compared to those in the voicing-promoting environment (38%). The effect of the voicing-impeding contexts was stronger, and very strong indeed was the incidence of completely devoiced tokens. When viewed in the voicing-promoting context, i.e. before a voiced sound, the plosives were fully voiced in 83% of the cases, and partially-devoiced in 17% of the cases, while the three possible realizations of the fricatives were equally frequent (Figure 2).

The present study is a continuation of the research into the phonetic voicing of all word-final voiced obstruents, with a desire to establish quantitative data with regard to the Voicing-Into-Constriction (VIC) interval.

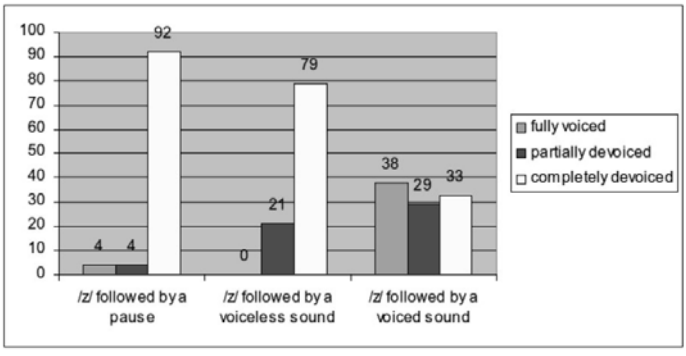

Figure 2: Voicing of word-final fricatives depending on the following context.

4. Experimental setup

The material was obtained from parts of widely available English TV programs listed in Table 1. The total time of recordings which were studied approached one hour and a quarter. The audio was extracted with the Adobe Audition 3.0 software, and the duration of the phonation segments in the speech of male and female speakers was measured on WASP. The figures were tabulated along with information on the preceding and the following contexts in which the studied sounds occurred altogether, VIC was measured in 340 sounds (198 fricatives and 142 stops).

Table 1: A list of language materials on which the present study is based.

No.	Title	Size	Duration
1	Ernie Almond chats with Jon Danzig HD	100 059 848	00:18:54.13
2	Holly Willoughby and pastry	18 032 024	00:01:42.06
3	Holly Willoughby 4-10-2011	59 603 820	00:05:37.26
4	Holly Willoughby Interviews Cheryl on Cheryl Cole's Night	31 584 060	00:02:59.01
5	Holly Willoughby talks about sleeping in a bath – Pea Shooting – This Morning	35 871 868	00:03:23.10
6	Jamie Bamber talks to Philip & Holly	31 091 398	00:05:52.14
7	Stephen Fry part 1	54 078 152	00:10:13.02
8	Stephen Fry part 2	50 432 710	00:09:31.23
9	Stephen Fry part 3	30 665 416	00:05:47.19
10	Stephen Fry part 4	40 061 638	00:07:34.05
11	Stephen Fry part 5	12 610 246	00:02:22.28
	Total	**464 091 180**	**1:13:54.47**

5. Results

Word-final obstruent voicing was studied from two basic angles to answer the questions specified in Section 2; for ease of reference we reproduce them again below:

1. Is VIC intrinsically related to the manner of articulation of the studied sound?
2. Does the preceding or following contexts exert any effect on the duration of VIC?

Let us now take them one by one.

5.1. Intrinsic voicing

The study has shown that the duration of the VIC interval is not related to manner of articulation, even at p=0.005 (Figure 3).

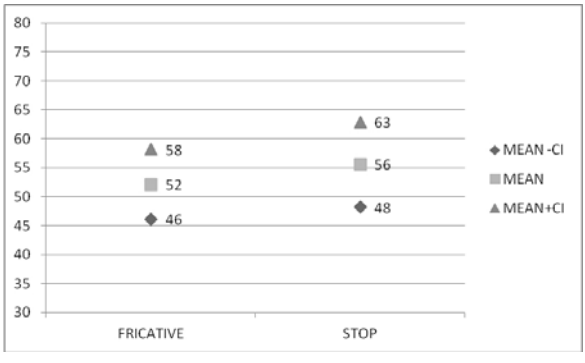

Figure 3: Intrinsic voicing in fricatives and stops at p=0.05. Here and henceforth, for each data point both upper (triangles) and lower (diamonds) confidence intervals are shown in the plots.

5.2. Extrinsic effects

The term 'extrinsic effect' refers to the preceding and following context of the studied obstruent.

5.2.1. Preceding context

The preceding context variable had three levels corresponding to the major sound classes: obstruent, sonorant and vowel. The study revealed that the preceding context exerted a non-significant effect on the dependent variable. Namely, at p=0.001, there was no significant difference between the effects exerted on the duration of VIC in the following obstruent by obstruents, sonorants and vowels (Figure 4).

5.2.2. Following context

The effect of the following context was first studied in relation to major classes. The study revealed a statistically significant difference between the stronger retroactive effect of vowels and sonorants vs. a weaker effect of voiced (V+) and voiceless (V-) obstruents and pauses (Figure 5).

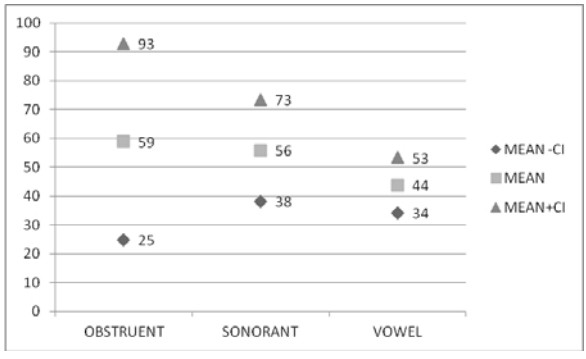

Figure 4: The effect of the preceding context on voicing.

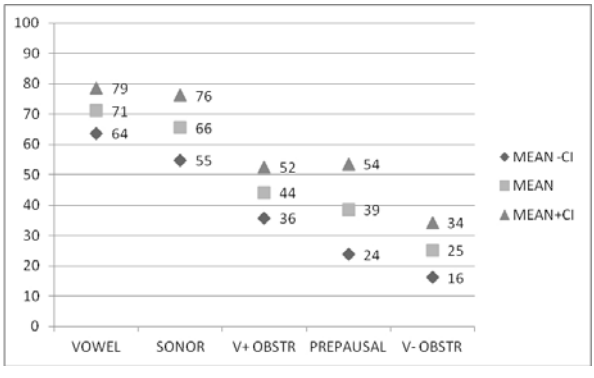

Figure 5: The effect of the following context.

Finally, the intrinsically voiced elements of the following context were pooled to obtain a bi-polar variable voiced-voiceless (Figure 6).

Figure 6 shows that the voicing of the sound following a word-final obstruent exerts a strong effect on the intrinsic voicing of the sound studied. In qualitative terms, the effect of the following context, when subcategorized into full voicing, voicelessness and

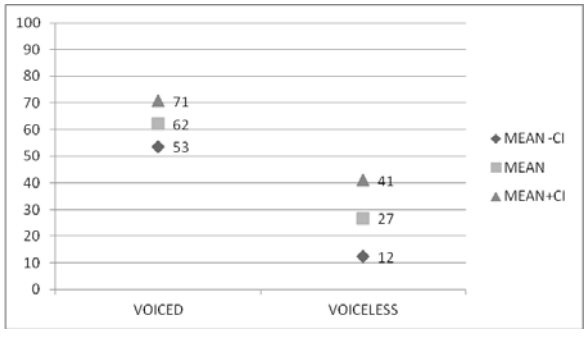

Figure 6: The effect of the following context.

intermediate values, yields the following results in the two studied contexts: voicing-impeding (Figure 7) and voicing-promoting (Figure 8).

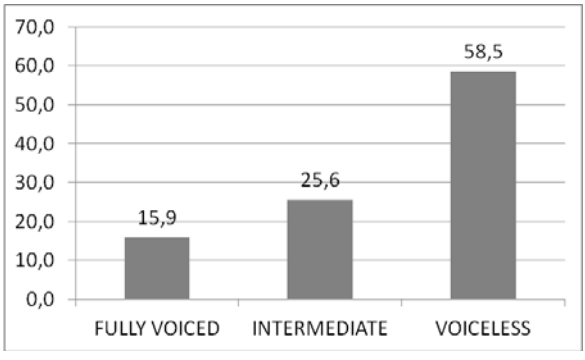

Figure 7: The voicing-impeding effect exerted by the following context on word-final obstruents (per cent).

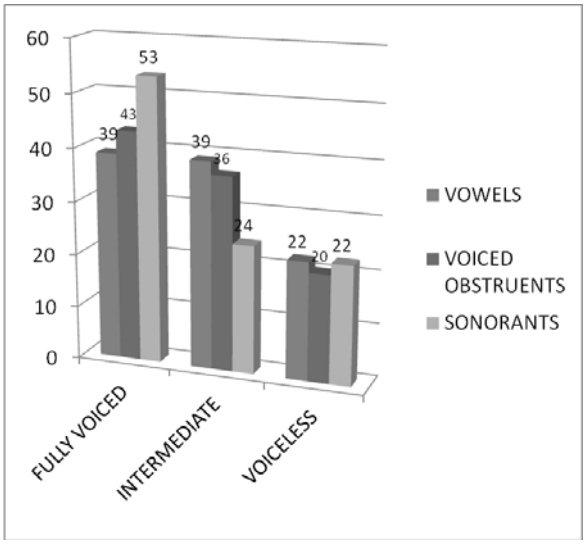

Figure 8: The voicing-promoting effect exerted by the following context (vowels, voiced obstruents, sonorants) on word-final obstruents (per cent).

6. Discussion

The present paper has revealed that neither the intrinsic characteristics of the obstruent, nor the preceding context had any significant effect on the duration of the VIC interval. However, the voicing of the following sound did have an effect on the duration of the VIC interval within the obstruent, with a mean of 62% of the duration of the constriction if the next word started with a voiced sound, and 27% of it if the next word

started with a voiceless sound. Yet both figures indicate that irrespective of the following context, the tendency to retain the voicing of the obstruent persisted.

A closer look on the distribution of full voicing vs. complete devoicing shows that in the voicing-impeding context more than a half of the lenis obstruents were fully devoiced, a quarter had an intermediate VIC value, and about 1/6 retained full voicing (Figure 7). A similar situation obtains in the voicing-promoting context, in which more than 40% of the lenis obstruents retained their full voicing despite the word-final position, less than 40% had intermediate values, and only 1/5 underwent full devoicing.

Our findings indicate that the phonetic voicing retention rules used in the context of teaching English phonetics should be relinquished in favour of a set of rules that take into account the following context and allow both more devoicing and more individual variability.

References

[1] Gonet, W. 2001. Obstruent Voicing in English and Polish. Pedagogical Perspective, *International Journal of English Studies* 1 (1), 73–92.

[2] Shockey, L. 2008. Personal communication.

[3] Davenport, M. and S. J. Hannahs. 1998. *Introducing Phonetics and Phonology*, London: Arnold.

[4] Ball, M. J. and J. Rahilly. 1999. *Phonetics The Science of Speech*, London: Arnold.

[5] Catford, J. C. 2001. *A Practical Introduction to Phonetics*, 2nd Edition. Oxford: Oxford University Press.

[6] Roach, P. 2000. *English Phonetics and Phonology. A Practical Course*. 3rd Edition. Cambridge: Cambridge University Press.

[7] Gimson, A. C. 2008. *Gimson's Pronunciation of English*. (Revised by Alan Cruttenden.) London: Arnold.

[8] Jassem, W. 1983. *The Phonology of Modern English*. PWN: Warszawa.

[9] Gonet, W. 2010. Dispelling the Myth of Word-Final Obstruent Voicing in English: New Facts and Pedagogical Implications. In E. Waniek-Klimczak, ed. *Issues in Accents of English 2*. Cambridge: Cambridge Scholars Publishing. 361–376.