Poznań Studies in Contemporary Linguistics 44(1), March 2008, pp. 37–60
© School of English, Adam Mickiewicz University, Poznań, Poland
doi:10.2478/v10010-008-0003-1

RP ENGLISH AND CASTILIAN SPANISH DIPHTHONGS REVISITED FROM THE BEATS-AND-BINDING PERSPECTIVE

MAŁGORZATA HAŁADEWICZ-GRZELAK
Opole University of Technology
haladewicz@gmail.com

ABSTRACT

This study sets out to determine the differences in the phonological representation of RP English and Castilian Spanish diphthongs. The analysis brings to light the discrepancies in the criteria for classifying a vocalic cluster as a diphthong. A functional explanation for the differences between Spanish and English complex nuclei in provided in Beats-and-Binding framework (the model developed by Dziubalska-Kołaczyk 1995, 2002a, *inter alia*). The discussion also shows that English and Spanish glides phonologically differ by their ability/inability to propagate a n→B binding. The conclusion is that "diphthong" is an *ad-hoc* notion which denominates completely different phonological entities in English and Spanish. What follows, it could be successfully supplanted by the functional term "complex beat".

KEYWORDS: Beats-and-bindings; Spanish and English diphthongs; glides; complex beats; binding propagation.

0. Introduction[1]

This study sets out to investigate phonological differences between diphthongs in RP English (henceforth referred to as "English") and Castilian Spanish ("Spanish" henceforth). The research reported in this paper supports the claim that the term "diphthong" denotes different phonological entities in the respective languages.[2] A functional explanation in Beats-and-Binding phonology is provided to explain these differences in terms of ability/inability to propagate a binding. My research has also shown that either the criteria for classifying a vocalic cluster as a diphthong should be unified, or the notion "diphthong" can be abolished altogether as superfluous.

---

[1] This article is an extended and upgraded version of a paper read at 35th Poznań Linguistic Meeting (Poznań 2005).

[2] Complex nuclei may be long vowels, or heavy syllables or multimoraic units, etc., depending on a theory, but any type of complex nuclei of differing vocalic quality is assigned the status of a diphthong.

There is a growing body of scholarship that focuses squarely on the study of diphthongs and hiatuses in English and Spanish.[3] Bogacka (2007) offers the most recent phonetic analysis of English diphthongs in the Beats-and-Binding perspective. Analyzing in particular the production of English diphthongs by Polish native speakers, she arrived at interesting juxtaposition of formant dynamics and vector length (the same reference for the exhaustive bibliography on the topic). Wolfram et al. (2004) provide a sociophonetic contrastive study of English and Spanish /ai/, Donegan (1985) is the reference for the analysis of vowel quality in Natural Linguistics. The idea of marking the most prominent element of a diphthong [+peak] and the less prominent element [−peak] which I use in the paper is taken from Awedyk (1974b).

Spanish vowel contact has also received a lot of scholarly attention. For example, Hualde and Prieto (2002) established that the acoustic difference between Spanish diphthongs and hiatuses lies in the mean duration values of the two vocoids. Cabré and Prieto (2004) provide a broad empirical study of the present-day lexical distribution of the rising diphthong/hiatus alternations in Central Catalan and represent the results in the form of OTed tableaux. They contend that "the tendency to contract vowels into rising diphthongs in Catalan can be regarded as a process closely guided by competing prosodic and analogical conditions, i.e., by forces which optimise prosodic structure and by forces which tend to keep identity relations" (Cabré and Prieto 2004: 115). Interestingly, the results also corroborated the irrelevance of syllable structure in the investigated issue: "[t]he presence of rising diphthongs in Catalan seems to be strongly controlled at the foot level rather than at the syllable level. Evidence that syllable structure does not play a substantial role in gliding comes from examples showing that the structure of the onset or the rhyme is irrelevant for this process" (Cabré and Prieto 2004: 124).

The research on Castilian Spanish hiatuses is expounded in Cabré and Prieto (2006). The authors investigated the distribution of hiatus and diphthongs in sequences of vocoids of rising sonority within the lexicon (excluding historical diphtongs). The study established main factors contributing to glide formation: "[g]eneral anti-diphthong environments are those in word-initial position: *v*[i.o]*la* 'viola', *d*[i.a]*blo* 'devil', *c*[i.a]*nuro* 'cyanide'. By contrast, glides are preferred in word-medial and word-final positions (e.g., *histor*[ja] 'history', *id*[jo]*ma* 'language'"(Cabré and Prieto 2006: 212). The distance of the vocoid sequence from the main word stress was found to be "another factor which conditions glide formation: the greater the distance, the greater the tendency to pronounce a diphthong" (Cabré and Prieto 2006: 215).

Aguilar (1999) provides further phonetic insights into the phonetics of Spanish glides and diphthongs. Her study had two objectives: " (a) The search for acoustic cues for hiatuses and diphthongs, in the temporal and frequency domain, first separately for each speech situation, and later, comparing both speech situations. (b) The observation of the behavior of vowel sequences with respect to phonetic reductions" (Aguilar 1999:

[3] Since this paper is dedicated mainly to English-speaking readers, naturally the part referencing the Spanish language is enhanced.

63). Among other things, her study showed that "the hiatus-diphthong distinction is acoustically signaled by changes in duration and in formant trajectories, especially of F2: hiatuses show a longer duration and a greater degree of curvature of the F2 trajectory than diphthongs" (Aguilar 1999: 72).

Hualde (e.g. 2004) offers an exhaustive review of phonology on "quasi-phonemes" in Spanish. With respect to glides, he concludes that

> Paradoxically, although all glides are contextual allophones of the high vowels in Spanish, there is a contrast between high vowel and glide in certain positions. There are surface contrasts because in a relatively small number of cases and in very limited contexts there is exceptional syllabification as hiatus of sequences that according to the general pattern of the language should be expected to be realized as diphthongs (the historical tendency to reduce hiatus to diphthong has been blocked). Exceptional hiatus sequences are not randomly distributed in the lexicon and, in fact, in some specific positions "exceptional" hiatus is the preferred configuration (in some varieties of the language)
> (Hualde 2004: 10)

Hualde (1997) also suggests that "the vowel phoneme /i/ has [i] as its single lexical allophone, but is realized as a glide [i] in contexts of syllable contraction. On the other hand, the glide phoneme [j] is consonantized to [y] or [ŷ] in syllable-initial position and is realized as a glide after a tautosyllabic consonant (semiconsonant or on-glide) and after a tautosyllabic vowel (semivowel or off-glide)" (Hualde 1997: 65f).

My analysis is phonological in nature. Without entering into a debate on the interrelation between phonology and phonetics, I intend to show that the analysis of phonotactic and phonological factors can yield vital conclusions, which of course can be further pursued in phonetic research contexts. In other words, I submit that phonetic cues are epistemologically as valid as phonological cues (in response to Prof. John Ohala's views, this is a case for "speculations"). We might propose mental representations basing on the computer image of sounds but equally, we might inspect how one sound behaves under the load of the contact with the adjacent sound. There is no basis for saying that the former analysis – only because it is machine mediated – is more relevant than the latter. From such a perspective, the phonetic image of the glides in English and Spanish might be the same or different. For me, the issue is epiphenomenal. The crucial thing is that phonologically they are different entities. In Spanish the glides form part of the beat (a syllable nucleus in traditional terminology) and in English they do not. In terms of the dynamic perspective this translates into the ability to draw stress by Spanish glides versus the lack of such ability in the English glides. This of course may or may not show in phonetic analysis.

The final issue to be explained at this point is the phonetic transcription adopted. English transcription is straightforwardly managed by the IPA symbols. Spanish phoneticians use either the symbols of *RFE* (phonetic alphabet of the *Revista de Filología Española*) or symbols of their own invention. In some works on Spanish phonetics the

IPA symbols appear together with *RFE* notations but they denote different sounds than the IPA specifications for the symbol in question. For example, [ɟ] in Spanish phonetic notation denotes a palatal voiced *affricate*, while according to the IPA it describes a palatal voiced *stop*, as for example in (Pol.) *giez*. Similarly, /ĵ/ in Spanish notation stands for a central voiced palatal fricative, while for example in English diachronic phonetic textbooks it describes the sound equivalent to /dʒ/ – a post-alveolar voiced affricate. Moreover, the very difference between the semi-consonantal and semivowel realization (/j: i̯, w: u̯/) seems to be a purely articulatory matter, like for example, English pre-fortis clipping. I decided to transcribe the Spanish palatal fricative by the IPA symbol /ʝ/. Similarly, the Spanish affricate allophone of /ʝ/ as, for example, in (Sp.) *Yo* 'I' is denoted by a digraph [dʑ] (as in Pol. *dzięcioł* 'woodchuck'), which I claim corresponds to the allophone in question. The notations are synopsized in Table 1. The left-hand side gives the symbols I adopted for the Spanish transcription and the right-hand side lists the symbols currently used for the sounds in question.

Table 1. The IPA transcription of some Spanish phonemes used in the paper.

The IPA symbols I suggest	The symbols currently used in Spanish phonetics research
/ʝ/ (palatal central fricative)	/J, y, ŷ, ĵ/; Stockwell and Bowen: [i̯, ĵ, ỳ, ž]
[dʑ] (palatal affricate)	(a) [ŷ]; (b) [d͡ʒ]; (c) [ɟ]

1. Theoretical underpinnings: Beats-and Binding phonology

The framework for this paper is Beats-and Binding phonology as developed by Dziubalska-Kołaczyk (e.g. 1995, 2002a, 2007). Beats-and-Binding Phonology (B&BP henceforth) is a syllable-less phonological model, grounded in Natural Phonology. To recall briefly, in Natural Phonology language is "a natural reflection of the needs, capacities and world of its users" (Donegan and Stampe 1979: 127). Furthermore, the underlying segments "are mental representations of sounds which are, at least in principle, pronounceable" (Stampe 1979: 35). What follows, the notion of explanatory adequacy in phonology cannot be theory internal but it must be based on phonetic facts and on the nature of human communication: "if a given utterance is naturally pronounceable as the result of a certain intention, then that intention is a natural perception of the utterance" (Donegan and Stampe 1979: 163).

The model has been subsequently developed by Dressler (1985, *inter alia*). Among vital contributions to the model, Dressler (e.g. 1985) follows Peirce in claiming that the more iconic is the sign, the more natural it is. Hence, the principles with in generative models are derived from UG, in Natural Phonology are assumed to ensure form the semiotic nature of communication.

As an explanatory model of Natural Linguistics,[4] B&B necessarily presupposes a functional bias (Dziubalska-Kołaczyk 2002a: 28). The model assumes two basic functions of language – the communicative and the cognitive function, the communicative function subsuming two main functions of phonology: perceptibility and pronounceability. The main properties of the functional explanation in Natural Phonology which are relevant to the thematic scope of this paper are the following:

– Goal conflict: mainly denoting the conflict between the tendency towards the speaker's ease of articulation (lenitions) and facilitating the processing of speech by the listener (fortitions) (Dziubalska-Kołaczyk 1995: 44).[5]
– Multifunctionality: one function can be carried out by several operations, and, conversely, one operation can serve several functions at the same time (Dziubalska-Kołaczyk 2002: 28).
– Hierarchy of functions: priority is given to the communicative and cognitive functions of a language. All other functions , such as for example, distinctivess, are subordinated to these two.
– Preferences: functional predictions have the form of predictions, which means that "natural linguistic universals, founded on functional and semiotic principles [...] have the form of preferences, and not of absolute statements" (Dziubalska-Kołaczyk 1995: 47).[6]
– The use of so-called external evidence on a par with internal evidence. The main areas of substantive evidence are: first and second language acquisition, sociophonology (phonostylistics), aphasia, loan word phonology and psychophonology (Dziubalska-Kołaczyk 1995: 48).

According to the B&BP tenets, the smallest functional unit of phonology is the beat:

> A beat is a unit rather than a measurement or device [...] and as such needs some referent in phonetic reality. It is expected to be better accessible than the mora, on one hand, and the syllable, on the other. Its functioning in phonology in relationships with other units of structure called non-beats (these relationships are called bindings) is expected to account better for the structure than the functioning of mora or syllable. […] In Beats-and-Binding Phonology, a beat is a *regularly recurring skeletal prosodic unit of phonological representation, of a size corresponding to that of a segment.*
>
> (Dziubalska-Kołaczyk 2002: 86 [italics in the original – MHG])

---

[4] For an exhaustive presentation of Natural Linguistics, see e.g. Donegan (1985); Donegan and Stampe (1987); Dressler (1985, 1999); Dziubalska-Kołaczyk (2002a, 2002b, and extensive references therein).

[5] For the interpretation of lenitions/fortitions in terms of effort management, cf. Kul (2007).

[6] The principle that phonetic laws are relative and not absolute goes back perhaps to Jespersen (1904). The laws of preference were clearly formulated by Zabrocki (1961 [1980]). The notion of "preference" is discussed in detail in Dziubalska-Kołaczyk (2002a, Chapter 4).

According to Dziubalska-Kołaczyk (2002: 100),"[r]hythmic preferences form the skeleton of phonological structures". The primary rhythm units are feet, the beats being their constituents. The universal preference is a trochaic pattern consisting of two beats, the first beat being preferably strong and the second weak. The preference for a binary foot pattern can be subsumed under the universal preference for binary paradigmatic and syntagmatic contrasts (Dziubalska-Kołaczyk 2002: 89 and references therein).

A beat ("B") is realized by a phoneme traditionally occupying the slot of the "syllable nucleus". Preferentially it is a vowel, but secondarily a consonant may acquire the function of beat, it is then marked as "N". A vowel qualifies better as a beat due to its greater saliency, based on its sonority value plus its articulatory openness. Accordingly, consonants possessing such features in a larger degree as more likely to become beats, for example [n̩] in English, /l̩/ and /r̩/ in Czech.

Beats (B) and non-beats (n) in a sequence are joined in a binary fashion by means of sonority-based bindings (Dziubalska-Kołaczyk 2002: 94). For example, in a sequence {BnB} there may be at most two bindings: a B←n binding, where a non- beat is bound to the preceding beat, plus a n→B binding, where a non-beat is bound to the following beat, i.e. {Bn+nB}. "A beat, however, may potentially stand alone while a non-beat must be bound to a beat. Thus, in the {BnB} sequence there may alternatively be one binding only, combined with a single beat, i.e. either {B+nB} or {Bn+B}" (Dziubalska-Kołaczyk 1995: 61). Language-specific preferences determine the existence of a binding in such sequences.

The two bindings differ in strength. "The n→B binding, i.e. the binding of a non-beat to the following beat (preferentially realized by a /CV/ sequence) is always stronger than the B←n binding, i.e., the binding of a non-beat to the preceding beat (preferentially realized by a /VC/ sequence; Dziubalska-Kołaczyk 2002: 94). It must be observed, though, that the prototypicality of n→B binding does not necessarily presuppose the prototypicality of CV syllables (cf. Dziubalska-Kołaczyk 1995, 2002 for the overview of literature on the topic). Such a strength preference is listener-friendly, i.e. it serves the function of optimization of perceptuality, since "it is generally the case that the most salient acoustic modulations in a syllable occur near the CV interface" (Ohala and Kawasaki 1984: 117, as quoted in Dziubalska-Kołaczyk 1995: 62). The prototypicality of the n→B binding is perceptually based, but even on Level 1 other processes can arise that override this function.

The B&B notations are as follows:

– "B" for a beat;
– "n" for a nonbeat;
– "N" for a consonantal beat;
– → for the binding between a nonbeat and the beat n→B;
– ← for the binding between a beat and a nonbeat B←n.

The B&B scenario for the structuring of phonology involves the four-layered structure. (Levels 1–3). Each level subsumes a particular type of preferences which are enumerated as follows (Dziubalska-Kołaczyk 1995: 72f):

- Level 0: These preferences constitute the most general level of phonology, the level of rhythmical preferences.
- Levels 1 and 2, which are called levels of the universal perceptual preferences (Dziubalska-Kołaczyk 1995: 60f). Level 1 is the level of underlying phonological binding preferences between beats and non-beats, i.e. how the beats combine with non-beats. It determines the existence and number of bindings in CV, VC, VCV, CVC or CVCV sequences, how bindings arise and combine.
- Level 2 is the level of phonotactic preferences based on the preferred sonority distance among vowels (V) and consonants (C), i.e. where the consonantal and vowel clustering takes place.
- Level 3: The level of articulatory preferences, realized mainly by means of assimilations and reductions (Dziubalska-Kołaczyk 1995: 72).

The diverging phonotactic preferences between languages are grounded on the level of rhythmic preferences – Level O. Since such preferences constitute the framework of phonological structure, they are of crucial importance for the phonological analysis. According to the B&B, the type of isochrony obtaining in a given language conditions the binding preferences of that language. The default unmarked rhythm is foot-timing. Particular languages diverge from it either in the direction of beat-timing or stress-timing (Diubalska-Kołaczyk 2002a: 100ff).

Level 1 is the level at which the languages adopt strategies against hiatus on a language-specific basis: non-beats act against hiatus and if the vowel contact is not broken by non-beats:

- B+B is reduced to one beat (short vowel), which is a change of structure on Level 1, usually through synaloepha or syneresis processes, e.g. Sp. *lo odian* ‘they hate it’ → /ˈloðjan/;
- B+B remains a two-beat unbounded sequence (Level 1). On Level 2, however, such a cluster may be phonemically represented either by a diphthong or a long vowel (English) or simply remain as it is, allowing hiatus. This is the case occurring in Spanish, when synaloepha or syneresis does not apply – no change of structure on Level 1, e.g. E *law*, Sp. *ley* ‘law’.

Level 1 is also where the “heaviness” of “syllables” is determined. In phonologically conditioned stress a diphthong or a long vowel tends to attract stress, because of their double-beat phonological representation. Moreover, a beat of the type B←n may also draw in stress, which implies that “heaviness” in the B&B model is conveyed by means of beats and binding contained within a binary foot or a phonological word boundary.

Both the number of beats and bindings are counted to determine the overall weight (Dziubalska-Kołaczyk 1995: 70f).

The notion of sonority in the B&B model is based on acoustic parameters, with the reservation that there is no transparent phonetic criterion for sonority ranking. This reservation is expressed in the fact that out of 4 levels of phonological structure within B&BP, only on one level (Level 2) is sonority the driving force, never acquiring the status of the Sonority Sequencing Principle, though: "At the level of phonological bindings (Level 1) beats are uniformly more sonorous than the nonbeats. In objective terms, it is the degree of modulation in several acoustic parameters (amplitude, periodicity, spectral shape, FO, cf. Ohala 1990a) that decides whether a {nB} binding is uniformly stronger than a {Bn} binding" (Dziubalska-Kołaczyk 1995: 64). It can be thus understood as a "default intrinsic property of phonological segment, i.e. a property belonging to the level of intention" (Dziubalska-Kołaczyk 1995: 64).

Level 2 then is defined by the optimal sonority distances, which for a given language determine the forms in which the clustering takes place[7].

> Phonotactics (governed by universal principles) accommodates itself language specifically to a given timing pattern. Those clusters become stable in a language which obey universal phonotactic preferences (Level 2) and thus survive due to the sonority distance between their members. They do not succumb to the overwhelming Level 1 preference for the two nonbeat-beat bindings. Thus […] consonants do not bind (there are no interconsonantal bindings on Level 1) but they do combine into clusters on a phonotactic preferences level (Level 2) due to the sonority distance between them.
>
> (Dziubalska-Kołaczyk 1995: 73)

Level 3 is reserved for the function of ease of articulation, in other words, for speaker-friendly preferences to achieve articulatorily easy phonotactic sequences. In contrast with Level 1 and Level 2 where contrast was an underlying principle, Level 3 is driven by similarity (proximity law). The universally favored combinations are those ruled by the principle of the least effort. Since speaker-friendly processes are directed at facilitating the articulation of segment sequences, Level 3 is characterized by assimilations, reductions and all the other kinds of lenitions. If registers were to be discussed, it might be claimed that casual speech evinces mainly the processes described on Level 3 while

[7] There is a further qualification that should be added to this schematic account of the B&B theory. Dziubalska-Kołaczyk et al. (2007) and Bertinetto et al. (2007) pursue the question of how universal and language specific constraints interact. The answer is sought in the degree of phonotactic complexity. Accordingly, intersegmental cohesion is defined as a degree of cohesion between adjacent segments and the optimal shape of a particular cluster is defined by referring to the Net Auditory Distance. To wit, "[t]he measure of markedness is the overall sonority, understood as a perceptual effect brought about to the ear by manner of articulation of sounds (MOA) as well as place of articulation (POA) and distance in voicing (Lx). In fact, rather than the overall sonority, it is better to refer to a net auditory distance to which all the three factors contribute (sonority, place of articulation and voicing)" (Dressler and Dziubalska-Kołaczyk 2006: 251). In this work I chose to rely on the older version (Dziubalska-Kołaczyk 2002), i.e. sondis.

careful speech – those of Levels 1 and 2. The arising conflicts between levels, i.e. hearer-friendly and speaker-friendly preferences, are resolved language-specifically by a primary tendency towards balance (Dziubalska-Kołaczyk 1995: 72f).

2. The problem: Descriptive generalizations

2.1. The palatal glide

In English, semivowels have a functionally consonantal nature, although phonetically they resemble vowels. For example, it is correct to say *a yacht* instead of **an yacht*, as would be expected before a vowel. Roach defined semivowels as "raised high vowels in consonantal function" (Roach 2005: 159). According to Kreidler (1989: 68), English glides differ from vowels in being [−syllabic] but like vowels, they are [−cons, +cont, +son, −sib]. /j/ and /w/ are thus the nonsyllabic equivalents of the vowels /i/ and /u/, respectively. In terms of contrastive features they both can be described as [+voice, +high] and /j/ is [+front, −round], while /w/ is [−front, +round] (Kreidler 1989: 68).

Spanish semivowels, however, are both phonetically and phonologically treated as vowels, although e.g. Guffey (2002) gives Spanish semivowels a sonority value of 2, which is higher than a vowel and equals the English value for that class of sounds in classical phonological paradigms. In other words, Spanish glides are a part of a traditional "rhyme". This approach is shared by most scholars investigating the language, for example, by Harris (1983), Franch and Blecua (2001), Nowikow (1992) and Navarro Tomas (1990). As far as I was able to verify, only Stockwell and Bowen (1965) treat Spanish glides as clustering with consonants. Accordingly, I assumed the nuclear phonological status of Spanish glides on an authoritarian basis.[8]

The main support for the rhymal taxonomy of Spanish glides comes from stress assignment analysis. For example, Harris (1983), analyzing stress patterns of the native vocabulary and on the correctness judgments of hypothetical realizations shows that Spanish glides form part of a branching nucleus. To wit, *teléfono* has stress on the antepenult, yet the antepenult stress is impossible if the penult is heavy. Accordingly, nonce lexemes of the type **teléboina* are illicit. The Spanish stress rule, as formalized within the tripartite syllable structure, is formulated by Harris as follows: "[a]ntepenultimate stress is incompatible if within the penult the vowel is preceded or followed by a glide, or followed by a consonant or 'antepenultimate stress is incompatible with a three segment penult unless the first two segments of the penult are both consonants'"(Harris 1983: 11). The division of the syllable into two parts, i.e. onset and rhyme enables collapsing of the reservations into one rule: "Antepenultimate stress is impossible if the penult contains *either* a nonnull coda *or* a branching nucleus" (Harris 1983: 13).

---

[8] Definitely, a contrastive phonetic study on English and Spanish glides is merited and these issues are left open for further research.

Spanish also has the palatal fricative /ʝ/ which is considered to be an allophone of /j/ by some phoneticians and which has no phonetic correspondent in English. In fact, the orthographic representation (<y>) fosters misunderstanding and mispronunciation, because in most languages, English included, this grapheme denotes /j/. It seems that the distribution of Spanish /ʝ/ can be phonotactically equalled with the distribution of English /j/. Using external evidence (loanwords) shows this parallel. For example, *yoga* in Spanish is realized with /ʝ/ with the possibility of word-initial fortition into [dʑ]. In English the /j/ appears. Further similarity is of phonotactic nature: neither English /j/ nor Spanish /ʝ/ can occur flanking a vowel to the right (*V /j/ or *V/ʝ/). The palatal semi-vowel as such exists in Spanish, though it is often classified as an allophone of /i/ (cf. Hualde 1997). In contrast with /ʝ/ it is orthographically represented by the letter <i> and it appears only in diphthongs and triphthongs.[9] The areas of divergence between English and Spanish palatal glide are synopsized in Table 2.

Table 2. A summary of the differences between the English and Spanish palatal glide.

English /j/	Spanish /j/
Phonologically a consonant	Phonologically a vowel
Cannot occur in the off-beat position	Can occur in the off-beat position
No phonological connection with any of the vowels	Considered to be a allophone of /i/
Phonotactically and phonologically more similar to Spanish /ʝ/	A part of the fortitive chain: /i/ → /j/ → /ʝ/ → [dʑ] /j/ cannot occur in absolute initial positions

2.2. The labio-velar glide

The sounds of a language are usually definable in terms of the intersection of several categories of articulation. These categories, although indispensable to properly define a given sound, are not on the same level of importance in determining how the sound will interact with the sounds in its vicinity. Sometimes one of the categories is given prominence, sometimes the other. This is particularly true of English and Spanish /w/'s. Both of them are cases of complex articulation – they are labial, velar and gliding. Yet, in defining the English phoneme the labial quality is more important, while in Spanish it is the velar one which counts as more prominent (Stockwell and Bowen 1965: 64f).

[9] Such a distribution entails the terminological dichotomy of "semi-consonant" versus "semivowel". Both of those terms are used in Spanish phonetic descriptions to denote the palatal semivowel /j/. The distinction is purely terminological (or perhaps articulatory as, e.g. the case of English pre-fortis clipping) although some researchers, for example Franch and Blecua (2001: 284f) claim that there is a phonetic difference between the two realizations. The division indicates the respective pre- or post-nuclear position of the sound constituting the syllabic margin. [j] as a semi-consonant is used in rising diphthongs in a pre-nuclear position. [i̯], classified as a semivowel, occurs in falling diphthongs in a post- nuclear position.

The difference becomes evident in assimilation processes occurring in the respective languages. For example, in Spanish the obligatory nasal allophone before /w/ is [ŋ], according to the pattern of POA in this language. It is true even of loanwords, such as *sandwich*, which is realized usually as [saŋˈwitʃe]. In English, the obligatory nasal in such contexts is /n/ but sometimes the articulation /m/ can occur, e.g. [ˈsæmwɪdʒ] (Stockwell and Borwen 1965: 65), which means that it is the labial, not velar quality of English /w/ that induces assimilation.[10]

Another example of the precedence of the velar quality of /w/ in Spanish is the confusion between /w/ and [ɣw] clusters. While it is acceptable within standard pronunciation to insert [ɣ] in words with /w/ in word-initial position – for example as in *huerta* [ˈɣwerta] 'orchard' – it also happens that in [ɣw] intervocalic clusters the spirant is especially weak and might even be dropped altogether, e.g. words of the type *agua* 'water' can be realized as [ˈawa]. Furthermore, [ɣ] epenthesis can also occur in the environment /n/ + <hue> . In such a case, /n/, according to the nasal assimilation pattern, obligatorily becomes [ŋ], e.g. *un huerto* 'allotment' can be pronounced as [uŋ gwerto] (Navarro 1990: 142).

It thus appears that the phonotactics of Spanish /j/ cannot be compared to /w/. Contrary to /j/, there is no consonantal version of /w/. Onsetless beats of the type *huerta* tend to epenthesize a homorganic fricative [ɣ]. In connected speech, there can occur [ˠ] at the onset of words starting with /w/ but such articulation does not supplant or suppress the /w/ proper: *huerta* 'vegetable garden', which is in careful speech realized as /werta/ can also be pronounced as [ˠwerta] (Navarro 1990). The optional strengthening is thus effectuated by word-initial epenthesis of [ɣ] and not by changing the phoneme as such.[11] Table 3 shows the summary of the discussed differences.

Table 3. The juxtaposition of the differences between the English and Spanish labio-velar glides.

English /w/	Spanish /w/
Phonologically a consonant	Phonologically a vowel
Cannot occur in the off-beat position	Can occur in the off-beat position
No phonological connection with any of the vowels	Considered to be a allophone of /u/
No phonological correspondent in Spanish	No fortitive chain: in the phonotactic position corresponding to Spanish /ʝ/ [ˠ] can optionally occur .

[10] Although the /n/ in words such as *sandwich* can be realized as a velar nasal in some varieties of English, such as (varieties of) Scottish English. It might be noted in passing that Scottish English, just as Spanish, can be classified as a beat timed rather than stress-timed language.

[11] It might be noted in passing that also in Zabrocki's analytical paradigm the /w/ is not equaled to /j/. It is located closer to consonants: k, p, t – c – s – m – n – l – r – h – j –w – i – u – e – o – a. (Zabrocki 1960 [1980: 56]). We might also recall here that the preference hierarchy as elaborated by Roach (2005: 159) is: palatal → labial-velar → labial-palatal→ velar.

2.3. Diphthongs

The question is insoluble and diphthongs remain in a kind of system limbo
O'Connor (1973: 138)

Let us start the subsection with the juxtaposition of exemple definitions of a diphthong in both languages:

English:

– A vowel sound produced by two adjacent vowels in the same syllable whose sounds blend together (ie, oy, ow).
(<mdk12.org/instruction/curriculum/reading/glossary.shtml>)
– A pair of vowels that are considered a single vowel for the purpose of phonemic distinction. One of the two vowels is more prominent than the other. In writing systems, diphthongs are sometimes written with one symbol, and sometimes with more than one symbol (for example, with a digraph).
(<mediasrv.ns.ac.yu/unicode/www.unicode.org/glossary/index.htm>)
– Sound which cannot be represented by a single symbol, usually between successive phonemes.
(<www.hitl.washington.edu/scivw/EVE/IV.Definitions.html>)
– A complex speech sound that starts at one vowel and moves toward another, as in the sound at the end of joy.
(<www.fonts.com/AboutFonts/Glossary/_glossary_D.htm>)
– A vowel sound, occupying a single syllable, during the articulation of which the tongue moves from one position to another, causing a continual change in vowel quality, as in the pronunciation of a in English late, during which the tongue moves from the position of (e) towards (i).
(<www.scnt01426.pwp.blueyonder.co.uk/Articles/Language/Glossary.htm>)
– A phoneme where the mouth glides from one vowel sound directly into another. The "long I" sound is a diphthong, but for reading instruction only the sounds /oi/ as in boy and /ou/ as in cow are taught as diphthongs.
(<jan.ucc.nau.edu/dsc27/Resources/Phonics%20Glossary.doc>)
– A vowel sound that starts near the articulatory position for one vowel and moves toward the position for another
(<wordnet.princeton.edu/perl/webwn>)

Spanish:

– (Diptongo) A sequence of two vowels in the same syllable. Either the first or the second vowel will be treated as a semivowel.
(<www.cus.cam.ac.uk/~cjp16/spanish/linggloss.htm>)

- "[A] diphthong is any combination of two vowels in a word in a situation of an open transition […]; it is formed by two phonemes /i/, /u/ plus any vowel. If the cited phonemes /i, /u/ occupy the prenuclear position, their realization is [j], [w]; if they occupy the postnuclear position [i̯], [u̯]" (Franch and Blecua 2001: 414).
- "A diphthong is a union of two vowels that are pronounced in the same syllable" (Artuñedo and Donson 2000: 34).
- "When the accent falls on /i/ or /u/ the diphthong disappears and the vowels pertain to different syllables" (Artuñedo and Donson 2000: 34).

As can be seen, English diphthongs are vowels, during the articulation of which the tongue glides from one position towards the position of another vowel – /ə/, /ɪ/, or /ʊ/ within a syllable (a beat in the framework of this paper). For example, in /aɪ/, the first element specifies the beginning and the other the direction of the movement of the tongue. Alternatively, an English diphthong can be also defined as a vowel, the quality of which changes during the articulation, due to the movement of the tongue. This is the differentiating criterion from a "pure" vowel, because "pure" vowels do not change quality during their articulation. According to Roach, "English diphthongs participate as units, parallel to long and short vowels in MP alterations. Thus /div[aɪ]n/ → /div[ɪ]nity/" (Roach : 138).

Spanish diphthongs, on the other hand, are a sequence of a vowel + semivowel and a semi-consonant + vowel, depending on the position within the nucleus. As such, they warrant a basic identity question: are Spanish diphthongs monophonemic or biphonemic? The prevailing theories can be sketched as follows (I retain the symbols as in the originals, for the more detailed overview of literature see Franch and Blecua 2001: 421ff; Hualde 2004):

(i) [j, i̯, w u̯] are the allophones or combinatorial variants of /i/ and /u/;
(ii) there exist two consonant phonemes /y/ and /w/, of which [j, i̯, w, u] are allophones;
(iii) all the varieties are the allophones of /i/;
(iv) [y-ŷ], [j] and [i̯] are grouped under the unique phoneme /i/ which would be a *semivowel*, since it can function both as a nucleus and a syllabic margin; the same taxonomy applies to the velar series;
(v) the description is simplified by eliminating *y*, since the opposition *y-i* is limited to the C-V environment, hence the description would be simpler by eliminating *y* and introducing (-) instead, which indicated the syllabic boundary.

In English the gliding movement of the tongue during the realization of a diphthong is not equivalent to the articulation of consonantal glides, in spite of the similarities in terminology (Sobkowiak 1997). What follows, both phonologically and phonetically, the first sounds of *your* or *why* are different from the second nuclear elements of *bind* or *cold*.[12] In contrast, Spanish words such as *agua* 'water', *miráis* 'you look', which are sequences GV and VG respectively, are classified as diphthongs.

The claim of the dubious diphthongal status of Spanish clusters of glide + vowel is further corroborated if we compare the trajectories of Spanish diphthongs (in Stockwell and Bowen 1965) with the trajectories for the representation of the English clusters "unrounded palatal semivowel + vowel" (e.g. /jɑː/, /jɔː/) and "labio-velar semi-vowel + vowel" (e.g. /wiː/, /wɒ/) as presented by Gimson (1970: 214, 216).[13] Then the similarities do appear. Such clusters in English are longer than diphthongs and the length indeed corresponds to the cognate Spanish diphthongs. Hence, the obvious articulatorily grounded conclusion is to deny the existence of diphthongs in Spanish.

Let us next compare the phonological behaviour of Spanish and English diphthongs. As might be inferred from the definitions of diphthongs presented *supra*, diphthongs in both languages form a heavy nucleus (complex beat). And this is where the similarities end. An English diphthong is an indivisible segment, which means that the diphthongal nuclei cannot be split by any phonological rule to form part of two separate nuclei (beats), for example, *ear* /ɪə/ cannot possibly become a dissyllabic (two-beat) sequence through the application of any phonological or morphonological rule. Gussmann (2002: 22) also observes that both phonotactically and phonologically, there is no difference between a long vowel and a diphthong in English.

The story is different in Spanish, though (cf. the last definition). For example, *enviar* 'to send' with a 'diphthongal' realization becomes *envías* 'you send', where the diphthong vanishes, giving rise to a hiatus. The process might be also illustrated by the Spanish parsing of the conjunctions *y* 'and' *o* 'or' (*e* is used instead of *y* when the following words starts with /i/, *u* is used instead of *o* when the following word starts with /o/). The pronunciation and parsing of *y* and *u* changes according to the phonotactic environment, e.g. *jugó y comió* 'she played and ate' [xu·ɣoi̯·ko·mio], *mirar y escuchar* 'to look and to listen' [miɾar·jes·kutʃaɾ] (Navarro 1990; Nowikow 1992: 29).

Recapitulating, in English the semivowels can cluster with consonants as the onset cluster, i.e. traditionally they do not contribute weight. For example, in *twilight* /tw-/ is an onset cluster. In Spanish the semivowels cluster with vowels, for example in *buey* 'ox' there is a triphthongal realization. In English diphthongs are indivisible, while in Spanish they can be divided by a word-internal boundary as in *rey* /rei̯/ 'king' → *reyes* /re·ʝes/. Moreover, in Spanish diphthongs can also arise due to resyllabification across words: *jugó y comió* 'she played and ate' can be realized as /xu·ɣoi̯·ko·mio/ with the conjunction *y* forming the diphthongal cluster with the last vowel or with the first vowel of the second word. They can also vanish when the accent shifts: *esquiar : esquías* 'to ski', *fío* 'I confide' : *fiaré* 'I will confide'.

[12] An alternative interpretation proposed by Hockett (1955) discussed and criticized in Roach (2005: 137). The reader can also find the critique of SPE which treats diphthongs as a sequence of tense vowels plus glide in Roach (2005: 92).

[13] It might be noted in passing at his point that also in RP English the status of /juː/ has been debated.

Table 4 shows a resumé of the problematic issues which were singled out during the discussion in the subsection.[14]

Table 4. Contrastive specification of English and Spanish diphthongs.

Spanish diphthongs	English diphthongs	English GC clusters
Phonetically: a cluster (G)V(G)	Phonetically: a vowel cluster V + V	Phonetically: a cluster G+V
Phonologically: a complex beat (in traditional terms: complex syllable nucleus)	Phonologically: a complex beat (in traditional terms: complex syllable nucleus)	Phonologically: the glide has a consonantal status – onbeat position, in traditional terms: syllable onset + simplex nucleus
Connection with stress assignment (e.g. *puerta – portal*)	Connection with stress assignment: phonologically: the same as long vowels	Stress assignment: count as "staple" CV sequence
The resulting complex beat can be split through stress rearrangement, e.g. *esquiar – esquío* 'to ski, I ski'	The resulting complex beat cannot be split through stress rearrangement	The resulting complex beat cannot be split through stress rearrangement[15]
The cluster can disappear through synaloepha (una agenda), syneresis (preeminente) or the fortition of the glide into /ʝ/ (*ley : leyes*).	The cluster cannot be split phonologically.	The cluster can disappear through affrication across word boundaries, e.g. *would you*, *hit you*.

2.4. Triphthongs

The existence of three vowels in a syllable is a triphthong. The Spanish definition affirms equally that: "Un triptongo es la unión de tres vocales en una misma sílaba. Ejemplo: *averiguáis*" (<http://es.wikipedia.org/wiki/Triptongo>). Interestingly, the

[14] By means of an interpolation it might be recognized that English and Spanish are by no means the only area where there are problems in applying the blanket term "diphthong". For example, KLV (1990) allow for two different representations of the French GV structures as in (Fr.) *ouate* , where the /w/ occupies once a slot of the onset and once, a branching nucleus (KLV 1990: 228). In the French coursebook *Khmer au Quotidien* (Filippi et al. 2003: 8), the authors mention that "il n'existe pas vraiment de diphtongues en français ['in fact there are not any diphthongs in French'].

[15] Cf. note 15 above. On the other hand, for example in Polish, the sequences glide + vowel always split in declension, e.g. (Pol.) *maj* (m→a← j) 'May' versus *ma.ju* (m→a j→u) 'May, loc. (Wiesław Awedyk, personal communication, April 2004).

given example is a sequence glide + vowel + glide: /-wajs/. What follows, in Spanish the most open vowel, which is usually the one articulated with the greatest articulatory energy, is the most prominent part of a triphthong, i.e. the nucleus [+peak]. Preceding the peak we find a semi-consonant and in a post-nuclear position there is a semivowel, both of which are [−peak]. A Spanish triphthong starts with an articulatory movement of increasing aperture and finishes with a decreasing one. Spanish triphthongs present a wide plethora of possible realizations, due to the lack of restrictions on diphthongal constituents, the lack of restrictions on word-final vowels, the lack of vowel reductions and accentual shifts (Franch and Blecua 2001). Triphthongs can also arise across words due to the restriction that a maximum of three vocalic elements can form a nucleus in Spanish.

Some illustrative examples are as follows:

(1a) *estudiáis* 'you (pl.) study'
(1b) *buey* 'ox'
(1c) *sentenciéis* 'you (pl) sentence'
(1d) *averiguáis* 'you (pl) verify'

English triphthongs, in turn, are very restricted and subject to further reductions. They can undergo diphthongization (smoothing, cf. Cruttenden 1994: 128) and, consequently, monophthongization (Cruttenden 1994: 129; Sobkowiak 1997). The structure of a lexeme is not relevant to their occurrence: they can occur both intra- and inter-morphemically. However, a linking may also arise across word boundaries. They can be described as being composed of 5 closing diphthongs with /ə/ added at the end as illustrated in the following examples (adapted from Sobkowiak 1996: 157–160).[16]

(2a) /eɪə/ *layer*, *convey-er* → /eːə/
(2b) /əʊə/ *coalition*, *grow-er* → /əː/, /ɜː/

/aɪə/ *tyre*, *suppli-er* → /aːə/	/aʊə/ *tower*, *allow-ance* → /aːə/
/ɔɪə/ *voyage*, *buoy-ant*	/ɔːə/ (rare)

A schematic representation according to the criteria in Awedyk (1974b) is shown in Figure 1.

In Spanish, the most prominent element is the central vowel, which is the most open one. In English, on the other hand, the middle /ɪ/ and /ʊ/ are the weakest elements and in connected speech they are reduced to a transitory glide or deleted (Cruttenden

[16] It is also debated whether words such as *fire* are monosyllabic, with triphthongs, or bisyllabic, and thus without triphthongs and the sequences in words such as *fire* and *power* can indeed lose the final bordering element, when they undergo monophthongisation.

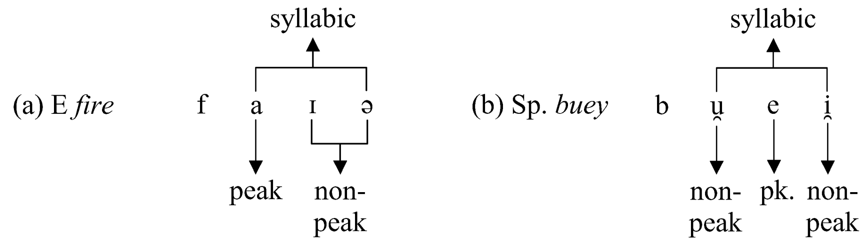


Figure 1. (a) The structure of an English triphthong (adapted from Awedyk 1974b: 84).
(b) The structure of a Spanish triphthong specifying the peak elements.

1994: 128). As a result, English triphthongs are usually reduced to diphthongs, which are transcribed above on the right of the glossed words in (2). Noteworthy is the fact that it is the central element that first of all undergoes elision. In Spanish, if the concatenation undergoes elision, it is always one of the bordering elements that is deleted, never the central element. Comparing the structures in terms of the constituent elements it could be concluded that /ɪ/ and /ʊ/, which are the only allowed central elements for English triphthongs, are in turn the only segments that cannot form a triphthongal peak in Spanish (/i/ and /u/ respectively). The structure of an English triphthong can thus be represented as a diphthongal glide which is followed by a schwa, "either as an inseparable part of the word, e.g. *Noah*, *fire* [...] or as a suffix (morpheme appended to the root), e.g. *greyer*, *player* [...], or, sometimes, as a separable element internal in a composite form, e.g. *nowadays* /naʊədeɪz/" (Cruttenden 1994: 128). To give a "pedestrian" illustration of the problem: English *way* is not a triphthong, while the similar sequence in Spanish as in *buey*, is.

3. Discussion in the Beats-and Binding framework

English semivowels, although phonetically they have features of vowels, phonologically behave as consonants, i.e. they do not form diphthongs (part of complex nuclei – complex beats). Spanish semivowels, on the other hand, are uniformly classified as part of a diphthong. For example, the /w/ of E *way* is classified as an onset, while in Spanish the same sound in *suerte* 'luck' constitutes the first element of a diphthong. It might also be noted in passing that English semivowels never occur in the off-beat position (traditional syllable coda). In terms of traditional syllable-based models English glides are the syllable onset.

I suggest to formalize this difference as the respective existence/lack of a binding between /j/ or /w/ and a following vowel, as in Figure 2.

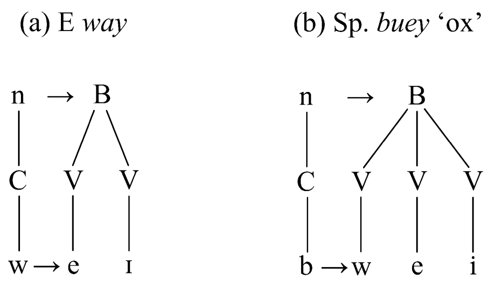


Figure 2. (a) The representation of E *way*. (b) The representation of Sp. *buey*.

The preferred beat representations in Spanish could be thus shown in the scheme in Figure 3.

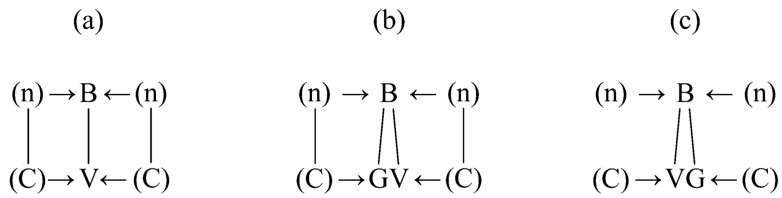


Figure 3. Preferred beat representations in Spanish in the B&B model.

What follows, I suggest taking into account the parameter of *binding propagation*, which is not contextually motivated but might be considered as a level 1 preference:

English:

– The glides can propagate/sustain an n→B binding: they can form CG clusters (e.g. a phonotactic chain CG→V.
– The glides cannot propagate /sustain B←n binding: they cannot occur in a phonotactic chain *V←GC or *V←G.[17]

[17] In an ongoing research project (cf. Haładewicz-Grzelak 2007), I motivate and formalize the claim that this phenomenon is part of a larger process affecting RP's most sonorous consonants: glides, /h/ and liquids: the inability to occur in the off-beat position, interpreted as the inability to resist lenitive aspects of B←n binding (weak position).

Spanish:

– The glides cannot propagate any binding
– They are drawn inside the beat to form a phonotactic chains of the type C→(G)V(G)(←C), the brackets denoting optional occurrence.[18]

Table 5. The contrastive specification of English and Spanish diphthongs in B&B.

Spanish diphthongs	English diphthongs	English GC clusters
No binding develops	No binding develops	The binding develops
G→|V|←G	(V)→|V|←(V)	C→V

As far as the term “diphthong” is concerned, I argue that with the view of the exposed ambiguities, the concept can be viewed as a sort of phonological hybrid: an ad-hoc language specific blanket term. Within the functional framework of the B&B model it is superfluous, just as the mora and the syllable. The representations I suggest can be managed by the following elements:

▯ – beat (B) ; □ – nonbeat, (n);

▮ – consonantal beat (N); ■ – a unit of sonority distance (sondis).

My extension involves introducing an additional notation, namely:

▯ – vocalic nonbeat (b).

A vocalic nonbeat denotes a segment in the beat area, which contributes its weight (traditional mora, in the B&B terms – quantity beat (β)) to the overall weigh count of the string. However, vocalic nonbeats do not develop a binding to a consonant, hence they are unstable and can be broken when the surrounding phonotactic conditions change. The binding goes directly from the consonant to the vowel and the vocalic nonbeat does not participate in it.

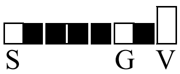
[18] To compare, Polish presents yet another option: the glides can propagate both n→B and B←n bindings, as in e.g. *miałki* /-aw-/ ‘powdered’, *kraj* /-aj/ ‘country’. Their consonantal status, judging only by phonological criteria, seems thus “stronger” than that of their English cognates. Furthermore, Polish glides provide validation for another claim adduced in this paper, namely positing a different phonological status of /w/ and /j/ formalized as more consonantal nature of /w/. To wit, in Polish the word-initial clusters GS with /w/ are possible, e,g. /wk-/ as in *łkać* ‘cry’ or /wg-/ *łgarz* ‘liar’. /j/ is not permissible in such phonotactic position.

The suggested representations rely on the scale as elaborated in Dziubalska-Kołaczyk (2002: 114). The B&B theory is constantly being upgraded (see Footnote 8 for the recent amendments to the theory – NAD) but for the argument reported in the paper the older version of the theory is sufficient; that is why I decided to use the version as it stood in 2002. The crucial difference from the traditional sonority counts (e.g. Guffey 2002) is the computational aspect. In B&B it is not the absolute value sonority values that are counted but the distances between particular classes of sounds. For example, the sondis between the plosive and the nasal is the same as between fricatives and semivowels (repeating the reservation that the presentation relies on the older version). What follows, the values are but a starting point for the subsequent computation. Table 6 presents the sonority count as used in this paper.

Table 6. The sonority scale (adapted from Dziubalska-Kołaczyk 2002: 114).

vowels	semivowels	liquids	nasals	fricatives	affricates	plosives
0	1	2	3	4	5	6

(1) The distance of a possible English cluster SGV is:



(2) The strategy adopted by Spanish is:

(3) The representation of an English diphthong would in turn look like:

In (3), the two vowels form an indivisible unit. Both the structures (2) and (3) form the so called complex beat, (2) representing the situation appearing in Spanish and (3) – the situation in English, which covers both the traditional diphthongs and long vowel nuclei.

In (1), the OSD in the part where a binding develops is equal to 1, which is the minimal value. In turn, in (2) the OSD is the largest possible, approximating the canonical CV shape. The process in case of Spanish can be schematically represented as CG ▸ ▸ V→C(GV).

In English the process seems to be the reverse. In a cluster stop + glide, the glide is subject to the influence of the preceding stop, whereupon the affrication arises iff the articulatory preferences on Level 3 are applicable. The process applies also across words (personal pronouns mainly): *tune* [tʃuːn] and *can't you* [kɑːn·tʃu]: C◂ ◂GV. In Spanish the affrication of a SG-cluster (e.g. as in *tiene* 'she has') is not possible.

The two strategies, although phonetically yielding different output, have the same function: to augment CV distance, further strengthening the n→B binding (cf. the "rich get richer" principle). Accordingly, it might be assumed that in Spanish a glide, which has the greatest level of "vowel-stickiness" is simply drawn by a vowel with no binding intervening. By means of such a process, two basic outcomes arise:

- the optimalization of sondis because any other sondis computation in a CV cluster is greater than Glide/Vowel (level 2 preference);
- moreover, it implies the cluster reduction of SG+V into the prototypical CV shape (Level 1 preference).

The "diphthongal" realization of an English cluster is shown in (3) (for the triphthongs, the representation would be analogous with three vocalic elements forming the beat). There is no phonological difference between the representation of a long vowel in English and an entity denominated as a diphthong.

The possibility for a functional explanation for the diverging status of glides in RP English and Castilian Spanish might also be sought on Level 0, that is the level of rhythmic preferences. The mechanical aspect of speech production cannot be denied. A sound wave is the propagation of energy through a medium, hence there are possibilities to explore stress/beat timing according to parameters used in mechanics. A stress-based language might have different tendencies for energy concatenation than a beat-based language. For example, in mechanics there is a so-called parameter of "stress concentrators" which is denoted, *nomen omen*, by the sigma sign. My suggestion, beyond what could be treated in the present article, would be to try to explain the ability /inability of a glide to develop a binding in this way.

4. Conclusion

In this article I have reviewed the status of glides and diphthongs in English and Spanish. The GV clusters are an ambiguous concatenation and it is only when phonological arguments are considered that the positional status of GV clusters may be discovered. My suggestion is to incorporate a parameter of *binding propagation* in the phonological description of glides. On that analysis, Spanish and English glides differ simply by the ability versus inability to develop a binding. English glides develop a n→B binding and the arising cluster has the phonological structure G→V. In Spanish, the glides cannot propagate/sustain any binding and they function as a part of a beat: (C)→GVG←(C).

Such a conclusion seems also to be supported by the fact that in English the glides can only develop the stronger, n→B binding. It means that English glides do not occur in a phonotactic position which would involve propagating B←n.

Furthermore, triphthongs and diphthongs in English and Spanish were reanalyzed and the existing research on Spanish diphthongs was critically reviewed. By means of conclusions the two definitions for diphthongs: the definition for English diphthongs and the definition for Spanish diphthongs were found incompatible. In English a diphthong is an indivisible entity, having phonologically the status the same as a long vowel. In Spanish, on the other hand, a diphthong can be created mechanically (as a morphonotactic cluster) and equally, can perish for morphonotactic reasons.

The discussion thus revealed that the concept of "diphthong" is language-specific, and similarly to "syllable" or "mora" it is a derived *ad hoc* term (i.e., where the definition precedes the phenomenon), hence it loses its relevance for contrastive studies. I would compare it to a phonological cenotaph: an empty place which misleads the research while the solution lies somewhere else. The suggested B&B formalization by means of a term "complex beat" could account for the areas of divergence without the need to manipulate the conceptual frames to include or exclude particular features of a vocalic cluster.

REFERENCES

Aguilar, L. 1999. "Hiatus and diphthong: Acoustic cues and speech situation differences". *Speech and Communication* 28. 57–74.

Alarcos, E.Ll. 1986. *Fonología española*. Madrid: Editorial Gredos.

Artuñedo, B. and C. Donson. 2000. *Curso de español para extranjeros* [Spanish course for foreigners]. *Cuaderno de ejercicios* [Exercise book]. Madrid: Ediciones SM.

Awedyk, W. 1974a. "The syllabic structures of English and Polish". *Papers and Studies in Contrastive Linguistics* 2. 89–95.

Awedyk, W. 1974b. "The distribution of non-syllabic phonemes in English and Polish". *Papers and Studies in Contrastive Linguistics* 2. 83–88.

Bańczerowski, J. (ed.). 1980. *Ludwik Zabrocki: U podstaw struktury i rozwoju języka* [Ludwik Zabrocki: At the foundation of language structure and development]. Warszawa: Państwowe Wydawnictwo Naukowe.

Bell, A. and J.B. Hooper (eds.). 1978. *Syllable and segments.* Amsterdam: North-Holland.

Bertinetto, P.P. and S. Scheuer and K. Dziubalska-Kołaczyk and M. Agonigi. 2007. "Intersegmental cohesion and syllable division in Polish". Paper read at ICPhS XVI Saarbrücken, Germany, 6–10 August 2007.

Bogacka, A. 2007. Repopulating vowel space: English diphthong production by Polish learners of English. [Unpublished PhD dissertation, Adam Mickiewicz University, Poznań.]

Cabré T. and P. Prieto. 2004. "Prosodic and analogical effects in lexical glide formation in Catalan". *Probus* 16. 113–150.

Cabré T. and P. Prieto. 2006. "Exceptional hiatuses in Spanish". In Colina, S. and F. Martínez-Gil (eds.), *Optimality-Theoretic advances in Spanish phonology.* Amsterdam: John Benjamins.

Cruttenden, A. 1994. *Gimson's pronunciation of English.* London: Edward Arnold.

Dinnsen, D.A. (ed.). 1979. *Current approaches to phonological theory*. Bloomington: Indiana University of Pennsylvania.

Donegan, P. 1985. *On the natural phonology of vowels*. New York: Garland.

Donegan, P. and D. Stampe. 1978. "The syllable in phonological and prosodic structure". In Bell, A. and J.B. Hooper (eds.). 24–34.

Donegan, P. and D. Stampe. 1979. "The study of Natural Phonology". In Dinnsen, D.A. (ed.). 126–173.

Dressler, W.U. 1985. *Morphonology: The dynamics of derivation*. Ann Arbor: Karoma.

Dressler, W.U., O.E. Pfeiffer and J.R. Rennison (eds.). *Phonologica 1988*. (Innsbrücker Beiträge zur Sprachwissenschaft.) Innsbruck: Institut für Sprachwissenschaft.

Dressler, W.U. 1999. "What is *Natural* in Natural Morphology (NM)?" *Prague Linguistic Circle Papers* 3. 135–144.

Dressler, W and K. Dziubalska-Kołaczyk. 2006. "Proposing morphonotactics". *Rivista di Linguistica* 18(2). 249–266.

Dziubalska-Kołaczyk, K. 1995. *Phonology without the syllable: A study in the natural framework*. Poznań: Motivex.

Dziubalska-Kołaczyk, K. 2002a. *Beats-and-Binding Phonology*. Frankfurt a.M.: Peter Lang.

Dziubalska-Kołaczyk, K. 2002b. "Challenges for Natural Linguistics in the twenty-first century: A personal view". *University of Hawaii's Working Papers in Linguistics* 23. 15–39. Also available at <http://ifa.amu.edu.pl/%7Edkasia/>.

Dziubalska-Kołaczyk K. and G. Krynicki. 2007. "Phonotactic preferences in Polish, English and German: Quantitative perspective". Paper presented at the 38th Poznań Linguistic Meeting,11–14 September 2007, Gniezno, Poland.

Filippi, J.M., H. Chan Vicheth and S. Sereyrat. 2003. *Khmer au quotidien.* [No indication of place]: Funan.

Franch, J.A. and J.M. Blecua. 2001. *Gramática española*. Barcelona: Editorial Ariel S.A.

Gimson, A.C. 1970. *An introduction to the pronunciation of English.* London: Arnold.

Guffey, K. 2002. *Spanish syllable structure.* Cambridge: University Press of America.

Gussmann, E. 2002. *Phonology: Analysis and theory.* Cambridge: Cambridge University Press.

Haładewicz-Grzelak, M. 2007. "Ludwik Zabrocki's structural phonetics and the substance of speech". Paper presented at the 38th Poznań Linguistic Meeting, 11–14 September 2007, Gniezno, Poland.

Harris, J. 1983. *Spanish syllable structure and stress in Spanish*. Cambridge: The MIT Press.

Hualde, J.I. 1997. "Spanish /i/ and related sounds: An exercise in phonemic analysis" . *Studies in the Linguistic Science* 27. 61–79.

Hualde, J.I. and M. Prieto. 2002. "On the diphthong/hiatus contrast in Spanish: Some experimental results". *Linguistics* 40(2): 217–234.

Hualde, J.I. 2004. "Quasi phonemic contrasts in Spanish". In: *WCCFL 23 Proceedings*.

Hualde, J. I. (forth.). "Stress removal and stress addition in Spanish". To appear in: *Journal of Portuguese Linguistics*.

Jespersen, O. 1904. *Lehrbuch der Phonetik*. Leipzig: Treubner.

Kaye, J., J. Lowenstramm and J.-R. Vergnaud. 1990. "Constituent structure and government in phonology". *Phonology* 7. 193–231.

Kreidler, C.W. 1989. *The pronunciation of English*. Oxford: Blackwell.

Kul, M. 2007. The principle of least effort within the hierarchy of linguistic preferences: External evidence from English. [Unpublished PhD dissertation, Adam Mickiewicz University, Poznań.]

Navarro T.T. 1990. *Manual de pronunciación española* [Manual of Spanish pronunciation]. Madrid: C.S.I.G.

Nowikow, W. 1992. *Fonetyka hiszpańska* [Spanish phonetics]. Warszawa: Wydawnictwo Naukowe PWN.

O'Connor, J.D. 1973. *Phonetics*. Middlesex: Penguin Books.

Quillis, A. and J. Fernandez. 1990. *Curso de fonética y fonología española* [A course in Spanish phonetics and phonology]. Madrid: Consejo Superior de Investigaciones Cientificas.

Roach, P. 2005. *English phonetics and phonology*. Cambridge: Cambridge University Press.

Sobkowiak, W. 1996. *English phonetics for Poles*. Poznań: Bene Nati.

Stampe, D. 1979. *A dissertation on Natural Phonology.* New York: Garland.

Stockwell, R. and J.D. Bowen. 1965. *The sounds of English and Spanish.* Chicago: The University of Chicago Press.

Wolfram, W., P. Carter and B. Morello. 2004. "Emerging Hispanic English: New dialect formation in the American South". *Journal of Sociolinguistics* 8(3). 339–358.

Zabrocki, L. 1960. "Zagadnienia fonetyki strukturalnej" [Problems of structural phonetics]. *Sesja Naukowa Komisji Filologicznej*. Poznań: Polskie Towarzystwo Przyjaciół Nauk. 165–185. Reprinted in: Bańczerowski, J. (ed.). 52–74.