This paper compares and contrasts the theories of Natural Phonology, Optimality Theory and Phonology as Human Behavior from diverse theoretical and methodological aspects including: the interaction between the opposing forces of markedness (the human factor) and faithfulness (the communication factor); the sentence-oriented versus sign-oriented approaches; and the concepts of naturalistic versus generative research paradigms. Despite these basic differences, similarities are also found in their shared functional basis which is discussed in the context of the natural phonological processes of Natural Phonology. I will further show how each theory views the notion of language universals. The concepts of combinatory phonology, phonotactics, and diachronic, developmental and clinical phonology will be discussed as measures of defining and determining the concept of language universals. The author maintains that biological, physiological, cognitive, psychological, sociological and other universals of human behavior are merely reflected in language rather than being specific “language universals” per se.

KEYWORDS: Natural Phonology; Optimality Theory; Phonology as Human Behavior.

1. Introduction

In this paper I will compare and contrast Natural Phonology (NP) (e.g. Donegan and Stampe 1979; Dressler 1996; Dressler et al. 1987; Dziubalska-Kołaczyk 2001, 2002; Dziubalska-Kołaczyk and Weckwerth 2003; Stampe 1972/1979) with Optimality Theory (OT) (e.g. Archangeli and Langendoen 1997; Kager 1999; McCarthy 2002, 2003,
The comparing and contrasting of the above theories are based on the following set of principles originally presented in Tobin (1990a: 68–69) for linguistic theories in general and reiterated upon in Tobin (2000b, 2007c) for phonology in particular. One of the most fundamental questions which must be addressed when discussing various linguistic or phonological approaches is: according to what theoretical and methodological principles can these different theories be compared and contrasted? And, even more fundamentally: what is the basis for determining these principles?

Every linguistic or phonological analysis is the direct result of a specific set of theoretical and methodological assumptions which are directly related to how the linguist or phonologist:

1. defines language and the sound system of language;
2. defines a linguistic or phonological problem;
3. determines the source, kind and amount of data to be selected and analyzed;
4. chooses a methodology to select and analyze the data;
5. evaluates, compares and contrasts analyses in light of all of the above.

These five criteria basically serve to describe how and what the particular linguist or phonologist views as the goal of linguistic or phonological research.

It is possible, therefore, to compare linguistic or phonological theories if they share a similar definition, or have a fundamentally similar view of language and the sound systems of language. This is certainly the case with NP and PHB because they are both “naturalistic” and “ecological” in their approach (as defined for NP by Gibbon 2007: 83–84, and clarified by Bogacka 2007: 99, and for PHB by Diver 1979, 1980, 1981). Saussure (1916/1972) has maintained that everything (in language) is an “opposition”, thus we can define “X by what it is opposed to in Y”, thus we can contrast NP and PHB as Naturalist theories in linguistics as opposed to generative OT based on the opposition that: Naturalism in linguistics has a history of opposition to abstractness, to Chomskyan generative linguistics, to formalist approaches.

Therefore, we may safely state that both NP and PHB are theories that form part of a larger “naturalist” paradigm which are opposed to the various syntactic, semantic and phonological theories found in the “formalist” generative paradigm (Gibbon 2007: 83), including OT. Advocates of this generative paradigm believe that formalization per se serves as the most basic theoretical and methodological criteria for the “scientific” explanation of linguistic phenomena (Tobin 2000).

1.1. Placing OT within the generative paradigm

If one accepts Kuhn’s (1962) thesis that scientific revolutions change the paradigms within which research is done within disciplines, then it is no surprise to the readers of
Comparing three phonological theories: NP, OT and PHB

this journal that there was a significant paradigm change in linguistics as a result of the Chomskyan revolution (from Chomsky 1957). OT represents one of the most interesting and innovative frameworks within the development of Generative Grammar which: (1) recognizes “deep” versus “surface” forms; (2) focuses on formal description; (3) searches for universal principles based on typology and language acquisition; and (4) aims to achieve economy in its analyses, among other paradigm specific principles.

Much of Kager’s and others’ work in OT stresses the differences between OT and other generative approaches to phonology and syntax. It should be remembered, however, that these differences are paradigm internal, and are, therefore, significant to other scholars working within the mainstream generative paradigm but are less significant to others who do not choose to work within the well-established paradigm of generative grammar such as advocates of NP and PHB.

It should be mentioned, however, that NP and PHB also differ regarding their affinities and similarities to various aspects of the formalism and the generative paradigm. Stampean NP is closer to generative grammar in general and OT in particular regarding the four criteria listed above:

(1) The underlying representations of NP may not be as abstract as those of OT, however, both NP and OT require that phonology should be capable of making correct predictions about any attempted utterance and do not limit themselves to the lexicon of the language under description. PHB, on the other hand, generally does not deal with abstract underlying representations and tends to limit itself to language-specific analyses. PHB also does not formally deal with language universals per se but recognizes that similar phonotactic and other favorings appear across languages and may be predictable to a certain extent.

(2) Although NP is often labeled as being not formal, this is not necessarily true, and NP has various formal requirements: rules versus prosody versus fortitions versus lenitions ordering, simultaneous application within blocs, implicational hierarchies, the similarity law, prosodic domains, distinctions between rules and processes, its restriction of perceptions to entities not dominated by contrary fortitions, etc.

(3) NP, like generative grammar and OT, searches for universals, not in the narrow sense but in the broad sense. How the original Stampean version of NP differs is that the universals are the processes actually used in speech production and perception (not simply metagrammatical statements of what is and what is allowed, or “preferred”). The language-specific, in NP, is the inhibitions of those processes. PHB, on the other hand, does not go from the meta-universal to the language specific inhibitions but first views language specific phonotactic preferences and then observes that these language specific favorings appear across languages.

(4) The concept of economy in generative grammar is achieved by fewer or simpler rules, whereas in the original Stampean version of NP, economy is achieved by
fewer inhibitions, and in OT, one may assume that economy is achieved by fewer faithfulness constraints. OT does break with the generative (Chomsky-Halle) paradigm, which assumes that all phonological alternations are learned, in proposing that phonological constraints are innate—so that some alternations result from ‘human limitations’. This makes OT closer in spirit to NP and OT. However, the concept of economy per se appears to be less of a major theoretical and methodological issue for PHB in general.

Kager (1990: 4) defines language or grammar for OT as a system of conflicting forces that are embodied by constraints which are universal but are ranked in a hierarchy which is not universal. Kager (1999: 1) further states that “[t]he central goal of linguistic theory is to shed light on the core of grammatical principles that is common to all languages”. The conflicting forces acting in OT are the constraints of “markedness” versus the constraints of “faithfulness” which are ranked differently and therefore may interact in a language-specific manner across languages in the generation of surface forms. My PHB-inspired interpretation of these opposed forces based on my naturalistic and ecological bias is that: (1) markedness represents the force of grammatical factors that exert pressure towards the prevalence of unmarked forms (more natural, less complex, easier to produce?) types of structure; (2) faithfulness, on the other hand, represents the force of combined grammatical factors that preserve lexical contrastiveness (clearer communicative distinctions and oppositions?). In the specific constraints they postulate and select, and in the particular hierarchical ranking of these constraints they adopt, OT phonologists generate surface phonological forms in a way that is unique among generative phonologists. Indeed, OT seems to have created a marriage combining functionalism and formalism in what appears to be an efficient and possibly even a well-motivated manner, making OT appear to be, in my opinion at least, to be close in spirit to other non-generative functional theories of phonology both formal and non-formal such as NP, PHB and Functional Phonology (Boersma 1998). For example, Stampe (1969) mentions conflicting processes which are universal but are resolved by ordering or suppression.

1.2. Contrasting generative OT with naturalistic NP and PHB

Natural theories, such as NP and PHB, unlike generative theories and OT, do not view formalisms as one of their central objects of interest but rather rely on the empirical centrality of external evidence related to biological, physiological, cognitive, psychological, sociological and other factors. In fact, the underlying assumption of both NP and PHB is that the acquisition and the phonotactic distribution of sounds in human languages can be directly compared with other instances of human behavior in general, and the need to learn how to control the musculature involved in performing fine motor movements in particular. NP also associates and finds the motivation for the acquisition
Comparing three phonological theories: NP, OT and PHB

and the phonotactic distribution of sounds in language to other domains of human life as well (defined by Dziubalska-Kołaczyk 2001 and clarified by Bogacka 2007: 99).

Furthermore, in addition to being “natural”, NP and PHB are also “ecological” in that they have been applied to language teaching and to the speech and language clinic (e.g. for NP: Dziubalska-Kołaczyk and Przedlacka 2005; Dziubalska-Kołaczyk and Weckwerth 2003; Grunwell 1987; Ingram 1990; Wrembel 2005; e.g. for PHB: Enbe and Tobin 2006, 2007; Fuks and Tobin 2008; Green and Tobin 2008a–b; Halpern and Tobin 2008; Reid 1991; Tobin 1995, 1997). OT, alternatively, may be viewed as being both formal and ecological in the sense that it, too, has branched out into developmental and clinical applications (e.g. Adam 2002; Adi-Bensaid and Bat-El 2004; Goad 1996). Moreover, to further justify my attempt in this article to compare and contrast NP, OT and PHB based on the fundamental similarities in their perspectives on language, I can point out additional developmental and clinical studies that have combined either NP and PHB (e.g. Połczyńska-Fiszer 2007) as well as OT and PHB (e.g. Adi-Been-Said 2006; Adi-Been-Said and Tubul-Lavi 2009; Ben-David 2001).

It should be mentioned, however, that when Gibbon (2007: 83–84) was originally referring to the “ecology” of linguistic theories, he was focusing on the central role of “technology” in “Natural Linguistics” in general and was referring to the level of “explanatory status” based on an “appropriate procedural formalization and appropriate computational implementation” in particular (which was compared and contrasted with the concept of explanation in NP by Bogacka 2007: 99–100). The concept of explanation in both NP and PHB centers on answering the question “WHY”. Although the concept of explanation in PHB and NP differ from the formal technology of OT, recent studies have applied and implemented PHB to a computer program based on lip-reading for the hearing-impaired (Schocken 2008; Schocken et al. 2008) making it “technological” as well.

It should be further noted, however, that when Diver (1975, 1980, 1981) originally used the word “ecological” with regard to his innovative approach to linguistics, he chose this term to describe the direct and integral connection between the forces of the human factor and the communication factor which Tobin (1990a, 1993, 1994, 1997) concisely summarized in the axiom that “human language represents a synergetic compromise in the fundamental struggle to achieve maximum communication with minimal effort”.

2. The Columbia School of Linguistics

PHB was developed by William Diver (1979) in an analysis of the non-random distribution of certain classes of initial consonant clusters in English, which he later expanded to explain the non-random combinations of vowels and consonants in English and in language in general (Diver 1993, 1995). PHB has been extended further to explain the combinatory phonology of consonant and vowel phonemes in a large number of different languages from several language families: e.g. Italian (Davis 1987), Hebrew (Tobin...
1990b–c), Urdu (Azim 1989, 1993, 1995, 1997; Hameed 2004; Jabeen 1993), Mewati (Fatihi 1987), Spanish (Flores 1997; Dekker and de Jonge 2006), English (Joue and Collier 1996), and Byelorussian (Dreer 2006). Diver’s original work in PHB has been expanded to explain the non-random distribution of initial consonant clusters in forty-two different languages representing nine diverse language families (Tobin 2000a). PHB has also been applied to the areas of developmental and clinical phonology for a wide variety of languages (e.g. English, Polish, Hindi–Urdu, Japanese, Finnish, Israeli Hebrew, Palestinian, Jordanian and Yemenite Arabic; e.g. Ali and Hameed 2007; Fatihi 2007; Miyakoda 2003, 2004a–b; Moore 1991, 1992; Tobin 1995, 1997; Tobin and Miyakoda 2004a–b, 2006); to the interfaces between phonology and morphology in diachronic inflectional and derivational morphology, and between phonology and the lexicon in the Hebrew triconsonantal (CCC) root system (e.g. Tobin 2004, 2006); as well as to a wide range of historical, psycholinguistic and sociolinguistic issues (e.g. Continik Morava et al. 2004; Davis et al. 2006; Reid et al. 2000; Tobin 1993, 1997, 2004).

PHB is part of a larger, functionalist, Saussurean sign-based theory of language formerly called Form-Content Analysis (FCA) and now referred to as the Columbia School (CS). FCA originally defined language as “a system of systems composed of various sub-systems revolving around the notion of the linguistic sign which are organized internally and systematically related to each other and used by human beings to communicate” (Tobin 1990a, 1993, 1994). This definition has been further simplified to: “language is seen as a symbolic tool whose structure is shaped both by its communicative function and by the characteristics of its users” (Tobin 2007a–b, 2008).

These definitions have both theoretical and methodological implications which highlight the differences between the sign-oriented CS versus the sentence-oriented generative paradigm (and even to a certain extent to other Natural Linguistic theories such as NP). In CS theory, for example:

1. The holistic Saussurean sign (signe linguistique) combining a signal (significant) and an invariant meaning (signifié) is the main unit of analysis, rather than the sentence; therefore CS theory does not recognize autonomous or separate levels of syntax and semantics (as well as the other well-established categories of “grammar”, such as subject, direct and indirect object, transitive, intransitive, etc.) which are widely accepted to be language universals;

2. CS theory does not distinguish between “deep” or “underlying” and “surface” forms concentrating on actual language use only; therefore, the theory does not employ derivations composed of formal rules or other constraints;

3. CS theory does not recognize a dichotomic distinction between grammar and lexicon but rather views them as a continuum;

4. CS theory concentrates on language-specific analyses rather than the more formal sentence-oriented concept of Universal Grammar (UG), as can be seen in a number
of anthologies that have primarily focused on sign-oriented CS (e.g. Andrews and Tobin 1996; Contini-Morava and Sussman Goldberg 1995; Contini-Morava and Tobin 2000).

2.1. Placing PHB in its historical framework

PHB may be viewed as part of the historical development of a larger twentieth century structural, functional, cognitive and naturalistic approach to linguistics. This tradition begins with Ferdinand de Saussure’s (1916/1972) concept of system and the dichotomies of langue and parole and phonetics and phonology based on a classification of sounds according to their articulatory and acoustic features. This fundamental dichotomy between the abstract code and its concrete realization based on distinctive features was further developed by Trubetzkoy (1939) and Jakobson (1941) of the communication-oriented Prague School (PS) (Tobin 1988). The strict communication factor adhered to by the PS was further supplemented by the introduction of the human factor to phonology through the concepts of “asymmetry” and “economy of effort in phonological change” by Martinet (1995). Martinet maintained that phonological systems are arranged asymmetrically and change in such a way that their non-random diachronic distribution reflects the search for equilibrium and harmony within the system as it is affected by the principle of least effort in human behavior. This principle of minimal effort postulated by Martinet implies that speakers strive for a minimal number of distinct phonemes which requires the least amount of effort to be produced and combined together in what Sampson (1980) reviewed in Tobin (1986) referred to as the “therapeutic view of sound change”.

It was Diver (1979), however, who has shown that a more complete theory of phonology has to take both the communication factor and the human factor into account. Diver maintains that there is a constant struggle between our need for maximum communication and our desire for minimal effort (referred to as “language synergy” in Tobin 1990a). The communication factor (requiring a large number of maximally distinct linguistic units demanding a great deal of effort) will be in conflict with the human factor (striving for minimal effort) resulting in a synergetic trade-off between the two. Therefore Diver (1979) extended Martinet’s (1955) more limited diachronic view of the human factor to that of a means of explaining the non-random distribution of phonemes in language in coordination with the needs of communication as originally established by Saussure (1917/1972) and the Prague School (discussed in Liberman 1991; and Tobin 1988, 1997).

3. Comparing NP, OT and CS: Where the triad meet

Despite the fundamental theoretical and methodological differences between NP, OT and CS briefly outlined above, they seem to share certain similarities in their outlook as well.
As we have previously stated, these similarities are embodied in the conflicting forces found in the constraints of markedness (the human factor) versus the constraints of faithfulness (the communication factor). Markedness for OT represents the pressure for the prevalence of unmarked (more natural, less complex easier?) types of structure. Faithfulness represents the preservation of lexical contrastiveness (more distinct and clear communicative distinctions and oppositions?). If we further analyze the sign-oriented definition(s) of language underlying CS provided above, we also find a conflict between the human factor (represented by: (1) intelligence; (2) economy of effort; (3) having large, but limited memories) and the communication factor (represented by the search for maximally distinct and clear-cut oppositions in linguistic units). This struggle between the human factor and the communication factor was summed up above in the following axiom which can now be adapted and compared to OT (and NP): Language reflects a constant struggle to achieve maximum communication (faithfulness constraints?) with minimal effort (markedness constraints?). This common perspective combining these interacting forces, factors, or constraints may explain why NP, OT and PHB share a fundamental interest in phonotactics, combinatory phonology and word and utterance structure, particularly regarding the non-random distribution of phonemes, the distribution of allophones, and phonological processes like neutralization, epenthesis, nasalization, etc.

In other words, all three theories seem to be interested in similar distributional phenomena, albeit from diverse perspectives. OT concentrates on deriving “surface” forms from “deep” forms using a variable set of formal hierarchical constraints on universal principles while both PHB and NP focus on external evidence to explain “WHY” we have the non-random distribution of phonemes found in specific utterances in various languages from the point of view of motivated principles linked to human behavior (PHB) and/or other domains of human life (NP). As a matter of fact, all three of these theories can be reproached in a similar manner by critics in search of hard and fast, or inviolate, formal universal phonological rules, for “juggling” the opposed forces of the human factor/markedness(?) with the communication factor/faithfulness(?), in an arbitrary or ad hoc way to achieve their language specific explanations of phonotactic distributions.

4. NP, OT and PHB: The shared functional base, or Funktionalismus über Alles!

Natural Phonology (Stampe 1972/1979) is one of the most predominantly used theories in developmental and clinical phonology. NP developed at a time when linguists began to question the way naturalness was handled within the generative framework. In developing his theory of natural phonology, Stampe started with the observation that the generative view of phonological rules as complexities that have to be acquired does not fit the facts of language acquisition. For example, devoicing of word-final obstruents occurs in German (or Russian), but not in English. Within generative grammar, German children are considered to “acquire” the devoicing rule. This, however, goes against what actually is observed, for, children do not “learn” devoicing, rather, they “never
learn” to pronounce “voiced” consonants in final position. Children whose native tongue is English commonly go through a stage where they devoice word-final consonants, but eventually “suppress” this tendency later on. It is apparent that in both cases, children start out with a tendency to devoice in final position. This fact could not be adequately accounted for within the generative framework. In NP, it is assumed that all children start out with the phonological process of devoicing, and as they learn to pronounce final voiced consonants, they “suppress” this tendency by eliminating the process of devoicing (in the case of German children, they do not “learn” to pronounce final voiced consonants; hence, the word-final obstruent is left devoiced).

This way of capturing phonological processes led to the integration of the concept of naturalness into phonological theory. Naturalness reflects what actually is observed in children’s speech. PHB claims that devoicing (using fewer sets of articulators; the human factor) generally takes place in final position because the demands of communication are lower in final position than they are in initial position (the communication factor). From the point of view of NP, the final devoicing in Russian, Polish, German, Dutch, Sanskrit, Korean, etc. are fortitive processes, e.g. in Dutch, Russian and Polish they apply before lenitions (such as voicing of obstruents before voiced obstruents). What is being communicated by final devoicing is the end of the syllable or the word. This section presents a selection of some of natural functional processes like those that Stampe described as they have been applied to developmental and clinical phonology (e.g. in Grunwell 1972; Ingram 1989; and Tobin 1997). These processes will be exemplified with Israeli Hebrew examples taken from the speech and hearing clinic and further elaborated upon and explained according to the principles of PHB:

A Functional processes influencing syllable structure

1. Final consonant deletion: CVC → CV (chronology: 2:0 → 3:2):
   word: sipûr → sipû ‘story’; gadól → gadó ‘big’
   syllable: taxtoním → tatoním ‘underwear’; jaldá → jadá ‘girl’
   syllable and word: maftéax → matéa ‘key’ (child with Down’s Syndrome)
   parpár → papá ‘butterfly’ (child with Dyspraxia)

   Explanation: Word final position has less communicative force; consonants require more articulatory control (are harder to make) than vowels.

2. Deletion of unstressed syllables (chronology: 2:0 → 4:0):
   one unstressed syllable: maká → ka ‘knock’; naxón → xon ‘right’; ambája → bájá ‘bath’; tsaláxat → láxat ‘plate’
   two unstressed syllables: tarnególet → gólet ‘chicken’; televízia → vízia ‘television’
Explanation: Stressed syllables give more communicative, perceptual and cognitive information than unstressed syllables; the more syllables in the word, the more effort it takes to pronounce it, in non-initially stressed words, there is less information in word initial position which usually has the greatest communicative force.

(3) Consonant cluster reduction: CC → C (chronology: 2:0 → 3:6-8):
initial: praxim → paxim ‘flowers’; bgadim → gadim ‘clothes’
medial: axálti/axálnu → axáti/axánu ‘I/we ate’ (child with Down’s Syndrome)
Explanation: A consonant cluster requires greater effort than a consonant–vowel sequence and may be reduced or replaced at the expense of maximum communication; in addition, coarticulation by near articulators is disfavored; phonemes of constriction (consonants) give clearer communicative distinctions than phonemes of aperture (vowels) – that is why there are more consonants than vowels in language – but they require more articulatory control (hence the ideal CV syllable).

(4) Reduplication (chronology: 2:0 → 2.5):
phoneme: kadúr → kúdu ‘ball’; dúbi → búbi ‘teddy-bear’
syllable: sáfta → táta ‘grandmother’; ejnát → nána ‘Eynat’ (girl’s name)
Explanation: The reduplication often comes as a means to avoid more difficult sound combinations and/or to maintain the number of syllables in the word; sequences of phonemes with the same articulators are disfavored unless their juxtaposition is, by virtue of some other factor, mutually beneficial. We also found that newly acquired sounds were often reduplicated as a means of practice or of hypercorrection in the clinical situation.

(5) Epenthesis: addition of segments (usually an unstressed vowel):
vowel: maftéax → mapatéa ‘key’ (+ stopping, + final consonant deletion)
(child with CP)
consonant: toxnit → toxsnit; toxná → toxsná ‘program’ (television/computer)
(two adult aphasics with Broca’s/Conduction Aphasia)
Explanation: The additional unstressed vowel/consonant often eases the transition within clusters and to more difficult consonants. The clusters then may be reduced at the expense of communication. Epenthesis, as well as the following Assimilation Processes in B (1) below, may also be explained by the factor: se-
Comparing three phonological theories: NP, OT and PHB

Environments of phonemes with the same or near articulators are disfavored unless their juxtaposition is, by virtue of some other factor, mutually beneficial.

B Assimilation processes (consonant/consonant–vowel harmony)

(1) Velar or nasal or labial, etc. assimilation (chronology: 2:0 → 2:8):
béged → gégé ‘clothing’; dubôn → dubô ‘little bear’; mángo → bámo ‘mango’

Explanation: A non-velar/nasal/labial sound changes to a velar/nasal/labial because of the influence of or the domination of a velar/nasal/labial sound which entails fewer articulatory gestures at the expense of maximum communication.

(2) Prevocalic voicing of consonants (chronology: 2:0 → 3:5)
kfiS → gviS ‘road’

Explanation: An unvoiced consonant (or consonant cluster) becomes voiced usually before a vowel: the speaker anticipates the control of two sets of articulators in what is usually a longer acoustic phonological segment. This is more important, even though additional articulators are disfavored, especially co-articulation by proximate articulators, in this case adding voice to the preceding consonant(s) in initial position with the higher communicative force.

(3) Devoicing of final consonants (chronology: 2:0 → 3:1):
bérez → bères ‘faucet’; od → ot ‘more’; jixtóv → itóf (+initial and medial consonant deletion; child with Dyspraxia)

Explanation: additional articulators are disfavored; voiced consonants become unvoiced in word-final (or medial) position: where the communicative force is less important or crucial, the speaker opts to activate one set of articulators rather than two.

C Substitution processes

(i) Processes reflecting the substitution of active articulators

(1) Fronting: back (non-apical) consonants are substituted by apical consonants usually preserving the same manner and voicing values (apicalization):
Explanation: the apex is the most flexible and easy to control of all the active articulators; the earliest and most frequent examples of the substitution of active articulators are fronting or apicalization which sharply reduces the number of communicative distinctions of the speaker.

(2) Backing: the back pronunciation of front sounds (usually consonants):

**Consonants:**
- mapít → maθít ‘napkin’;
- kos → koS ‘cup’
- batsék → axé (+ initial/final consonant deletion; child with Dyspraxia)

**Vowels:**
- éjfo → ófo ‘where’;
- lexem → laSem ‘bread’ (child with Down’s Syndrome)
- ani → aná ‘I’ (child with Cerebral Palsy)

Explanation: a (possibly idiosyncratic) later, less frequent process where the dorsum (or other back articulators) replaces the apex (or other front articulators); often found in children who have difficulty controlling the musculature of the apex (for organic or other reasons) and/or try to reduce the number of communicative distinctions made by the apex (or other front articulators) especially as a result of earlier fronting or other processes.

(ii) Processes reflecting the substitution of turbulence and/or airflow

(1) Stopping: fricatives/affricates are replaced by stops: variable chronology depending on sounds and language (chronology 2:0 → 5:0+):
- sus → tut ‘horse’;
- xatulá → katulá;
- tsav → tav ‘turtle’ (deaffrication)

Explanation: maximum constriction is favored particularly when mobile (stop) phonemes of constriction are easier to control than stabile (fricative) phonemes of less constriction which require greater control of the musculature to create and maintain a small aperture for a stronger turbulent airflow: the most frequent manner substitution for children.

(2) Gliding of liquids: l/r → j/w (variable chronology 2:0 → 5:0+):
- delet → dejet ‘door’;
- haláx → hajáx ‘he went’;
- mazlég → mazwég ‘fork’

Explanation: maximum aperture is favored: substitution of a lower to a higher degree of aperture (from consonants to semi-vowels) which may also require less articulatory control.
Comparing three phonological theories: NP, OT and PHB

(3) Glottal replacement: oral consonants are replaced by glottal stop:

varòd → vaʔòd ‘pink’; efrònòt → efʔonòt ‘pencils’; agalá → agaʔá ‘wagon’
ve-lo → ve-ʔo ‘and-no’; marák → ḥaʔák ‘soup’ (child with Down’s Syndrome)

Explanation: additional articulators are disfavored; a glottal stop replaces an
intervocalic consonant or a consonant in syllable-final position; articulatory
control of one set of articulators rather than two in an appropriate phonetic en-
vironment.

Thus we can present NP’s functional processes as a bridge uniting the naturalistic and
ecological approaches to a WHY-oriented explanation associated with NP and PHB and
therefore show the stronger link between NP and PHB in order to highlight their theo-
retical and methodological differences with the formal, generative-oriented OT.

5. Linking the theories together

However, to further illustrate the possible similarities and connections between NP and
PHB to OT, I will now list some of the major principles derived from the language spe-
cific analyses of PHB annotated with examples of the above functional processes
coined by NP which can also be directly associated with the markedness and faithfulness
constraints of OT:

(1) Additional articulators are disfavored, e.g. devoicing, nasalization, glottal re-
placement (human factor/markedness?).

(2) Coarticulation by near articulators is disfavored, e.g. consonant cluster reduc-
tion, deletions and substitutions (human factor/markedness?).

(3) Coarticulation by the same articulators/phoneme is even more highly disfa-
vored, unless their juxtaposition is, by virtue of some other factor, mutually
beneficial (cf. 13): e.g. reduplication, assimilatory processes, deletions, substi-
tutions (human factor/markedness?).

(4) Different word, stem or root positions have different communicative force
(higher communicative force in initial position, lower communicative force in
final position) and thus affect the favoring and disfavoring of different articula-
tory gestures, the number of articulators, the number and type of acoustic fea-
tures and the distribution of phonemes, e.g. prevocalic voicing of initial conso-
nants, devoicing of final consonants, deletion of final consonants, deletion of
unstressed or reduced syllables, initial epenthesis (human factor/markedness? +
communication factor/faithfulness?).
(5) Apical articulations are favored in general and in final position (lower communication force) in particular, e.g. fronting, apicalization (human factor/markedness?).

(6) Visual articulations are favored particularly in word/stem/root-initial position (higher communicative force), e.g. labialization (communication factor/faithfulness?).

(7) Explosive (stop) phonemes are favored in initial position, e.g. stopping/affrication (human factor/markedness?).

(8) Turbulent (fricative) phonemes are favored in final position, e.g. deletions, substitutions (human factor/markedness?).

(9) Transitions from one distinct constriction to another within a single phoneme (affricates) are disfavored (deaffrication) (human factor/markedness?).

(10) Consonant clusters are restricted concerning different articulatory and acoustic features (human factor/markedness?).

(11) Among constrictions, maximal constriction is favored, e.g. stopping (human factor/markedness?).

(12) Among apertures, maximal aperture is favored, e.g. vocalization, gliding (human factor/markedness?).

(13) Sequences of phonemes with the same articulators are disfavored unless their juxtaposition is, by virtue of some other factor, mutually beneficial (cf. principle 3 above), e.g. reduplication, coalescence, assimilatory processes: velarization, nasalization, labialization, prevocalic voicing (human factor/markedness?).

The following six principles were obtained from the developmental and clinical research performed within the theory of PHB following the natural processes established by NP which are directly related to the synergetic interaction of the human factor (markedness constraints of OT?) and the communication factor (faithfulness constraints of OT?):

(14) The preservation of as many distinctive features as possible (usually 2 out of 3) in substitution processes which require more effort than deletion processes (communication factor/faithfulness?).

(15) The preservation of as many communicative oppositions as possible in the original word (e.g. the number of phonemes per word) in substitution processes
Comparing three phonological theories: NP, OT and PHB

which require more effort than deletion processes (communication factor/faithfulness?).

(16) The use of a readily available phoneme already found in the speaker’s repertoire in accordance with the immediate phonetic environment in substitution processes which require more effort than deletion processes (human factor/markedness?).

(17) The preservation of the original phonetic structure of the word in deletion processes not involving syllable reduction and in reduplication (communication factor/faithfulness?).

(18) The preservation of the stressed syllable bearing the most communication information, if the original structure of the word is reduced by the deletion of syllables (human factor/markedness? + communication factor/faithfulness?).

(19) If the original structure of the word is enlarged by epenthesis, the epenthesis makes the transition to or between more difficult sounds easier (human factor/markedness?).

These principles empirically support the following conclusion underlying the theory of PHB: Developmental and clinical speech errors may be viewed as an extreme version of the mini–max struggle: there is less than maximum communication because of either extreme minimal effort or a lack of control over the articulatory tract or mechanisms. Greater effort will be exerted in order to achieve more efficient or better communication through clinical intervention. Therefore it may be assumed that the shared functional orientation of NP, OT and PHB allows them to overlap in the specific phonotactic issues upon which they focus and apply to developmental and clinical phonology, albeit to achieve different theoretical goals through diverse methodological means.

6. NP versus PHB and language universals

According to Dziubalska-Kołaczyk (2007: 71), the convenor of the special workshop on Natural Phonology – “Universal principles for the study of language (insiders meet outsiders)”:

Ever since Jakobson (or, shall we say, Plato?) linguists have been searching for universals (e.g. Maddieson 2006, 2007). Their views on the role of universals in language and linguistics have varied widely, though. In this session we want to concentrate on the question of the existence of universal principles for the study of language. NP (...) has always advocated the holistic view on lan-
guage, both in the sense of analyzing language structures (against “splendid isola
tion” of any structure) as well as in the sense of seeing language as part of
the universe. The latter means that the same principles of explanation apply to
language and to other aspects of life, and thus they are derivable from the most
general laws of human interaction with nature. In NP the principles are cogni
tive, phonetic, psychological, sociological, etc. They lead to the establishment
of linguistic preferences which guide the explanation of language-specific
structures. Apart from Natural Phonology, many other theories refer to univer-
Sals. Do they, however, look for universal principles of explanation? This is
the core of the planned debate between the NP insiders and the outsiders wish-
ing to take part.

As an answer to this question, I can state that both the naturalist theories of NP and
PHB differ from the formal generative approach of OT, as discussed above. NP and
PHB unabashedly advocate a holistic view of language both theoretically and method-
ologically. For PHB, this holistic approach to language is based on the following two
tenets: (a) all levels of language represent and embody the same mini–max principle re-
garding the compromise between the human and the communication factors; and (b) the
same principles found for language can be found in and directly supported by other as-
pects of human behavior that are extra-linguistic. Regarding which aspects of human
behavior and how they are related – for phonetics and phonology at least – can be asso-
ciated with the successful control of musculature needed to perform activities that are
related to fine motor movement. However, advocates of the theories of CS and PHB
primarily focus their attention on language-specific analyses only to determine which
aspects of human biological, physiological, cognitive, psychological, sociological,
pragmatic and other aspects connected to our behavior may be directly or indirectly re-
flected in and influence individual languages as a fundamental and integral means to
support our explanations of language phenomena. Thus, they may differ from followers
of NP (and OT) who assume a priori that there are language universals which underlie
and may determine in advance the way all languages work. However, both NP and PHB
attribute universals of language to universal characteristics of human production, per-
ception, and cognitive abilities which are not specific to some special faculté de lan-
gage. Thus, it is basically a question of which is the horse and which is the cart and
which one is put before the other? And/or who is the chicken and who is the egg and
who came first? Not all the postulated language universals appear in every language,
nor do languages represent the same shared universals in identical or similar ways.
Maddieson (2006, 2007) also points out that “typological frequency” does not indicate
“phonetic motivation” – which he views as being complex, diverse and random. There-
fore the potential role of “language” and phonological universals and the study of lan-
guage and phonology is still an empirical question requiring further research by diverse
theoretical and methodological approaches.
7. Concluding remarks

In this article, the theories of NP, OT and PHB have been compared and contrasted within the larger framework of phonological approaches within different linguistic paradigms. At this point of linguistic research, it is impossible to determine definitively whether one’s phonological or linguistic theory really reflects how language works in the human mind. Perhaps, when we know more about the human brain, we will be able to see which (if any!) phonological or linguistic theory has come close to capturing the “psychological reality” behind language. In the meantime, we can only continue to develop established theories and create new theories and compare and contrast these different theories by testing them against language data within and across languages. Perhaps, in this way, we may begin to reach an understanding of the psychological reality of phonological and linguistic theories, if such a thing is possible.

No theory has a patent on truth, and we all still have a long way to go. Moreover, there is neither a single truth nor a solitary approach to language and phonological and analysis in the first place. The more we study language and languages, the more challenging it becomes because we realize how little we actually know, and how complex the object of our research is. It should not be surprising, therefore, that phonology and linguistics are terribly splintered and fragmented disciplines. Many of us have placed ourselves, or have been placed by others, on what seem to be “theoretical bandwagons”, waving a lot of theoretical and methodological banners to and fro at each other.

One of the results of all this “flag-waving” is that we often seem to forget that there is more than one way “to skin a cat” – and do phonological and linguistic analysis. It should be important to phonologists and linguists to learn as much as they can about the various alternative approaches of analyzing the sound and other systems of language – if only to learn to discover and appreciate the similarities and understand and respect the differences.

REFERENCES

Adam, G. 2002. From variable to optimal grammar: Evidence from language acquisition and language change. (Unpublished PhD dissertation, Tel-Aviv University.)


Dreer, I. 2006. “Phonology as human behavior: A combinatorial phonology of Byelorussian”. In: Reid et al. (eds.). 107–130.

Comparing three phonological theories: NP, OT and PHB


Joue, G. and N. Collier. 2006. “Functional motivations for the sound patterns of English non-lexical interjections”. In: Reid et al. (eds.). 143–162.


Comparing three phonological theories: NP, OT and PHB


Tobin, Y. 2006. “Phonology as human behavior: Inflectional systems in English”. In: Davis, J. et al. (eds.). 63–86.


Trubetzkoy, N. 1939. *Grundzüge der Phonologie. (Travaux du Cercle Linguistique de Prague 7.)*


Address correspondence to:
Yishai Tobin
Department of Foreign Literatures and Linguistics
Ben-Gurion University of the Negev
P.O. Box 653
Be’er Sheva 84105
Israel
yishai@bgu.ac.il