Optimal Phonology: The Very Idea

Zhenjun Song

School of English Studies, Shanghai International Studies University, China

Abstract—The epistemology, theoretical constructs and production mechanism in optimality-theoretic approaches to phonology are examined and critiqued in terms of the phenomenology of speech sounds. An understanding has been arrived at that phonological optimality is construed too narrowly. A linguistic form is optimal in that it invokes in the hearer the meaning intended by the speaker. Its occurrences in different usage events form a cloud, in which phonological abstractions take place. Phonology is thus an inventory of both concrete and abstract exemplars, with the latter immanent in the former. Due to its emphasis on constructions, Cognitive Grammar provides the means for the flesh-out of a comprehensive model for optimality-theoretical approach to phonology.

Index Terms—phonology, optimality, OT, exemplar theory, cognitive grammar

I. INTRODUCTION

Anything in this world is subject to competing forces of various kinds, whether in its past, present, or future. The way of its being can be nothing if not fit for its survival. The concept "optimality", fitness or suitability of the one elected from many, therefore, captures the very essence of its phenomenology. An investigation of it in terms of optimality, the balancing of competing forces acting upon it, is consequently a very promising line of pursuit.

An object of scientific interest is the result of a long process of adaptation and selection, which usually spans over generations or even millennia. Those who wonder at and speculate about it are the human observers, who, because of the transient nature of their lives, have no chances to witness the whole process of its becoming. For its explication and prediction, one can only resorts to model-building to simulate the path of its evolution. To the effect that a model induces a result that is isomorphic with the object in question, it is considered to be a successful one: it explains the object or phenomenon and is capable of making predictions about its future development.

Model-building, however, is ridden with the investigator's epistemological bent. The kind of ontological status he attributes to a phenomenon under investigation is of paramount importance for the efficacy of a model. Due to the complex nature of most phenomena and the intricacy needed for its exposition, the ideology behind a model can easily fall out of the purview of a novice or even a professional sometimes. The result is usually a blind adherence to a biased model and its unjustified propagation, which in turn does more bad than good to healthy theory developments.

Optimality, due to its aforementioned potentials for explanation and prediction, has been introduced into the study of phonology by Prince & Smolensky (1993) in the early 1990s. The ensuing years is a dramatic period in which their approach has conquered and expanded. Few people have questioned the nature of the phonology as envisaged by their optimality-theoretic model. In view of the criticisms it has received from evolutionary phonology (Blevins, 1997) and its general apathy to the symbol grounding problem (Harnad, 1990), we deem it necessary to reexamine some of the key constructs in this model to put phonology and phonological optimality into a broader perspective.

II. METHODOLOGY

This is an argumentative thesis aiming to pave the road for a cognitive approach to phonology. Since its major concern is how speech sounds achieve "optimality" for effective verbal communication, the theoretical constructs in extant optimality-theoretic models have to be scrutinized for their validity and the oversights of critical aspects involved in linguistic behaviors in these models have to be pinpointed before a new proposal can be made. Our method is then one of presentations and critiques.

The key concepts and ideas in three models, classic Optimality Theory (OT) (Prince et al, 1993; Kager, 1999), functional OT (Boersma, 1999), and the Combined Model (van de Weijer, 2012), will be presented, analyzed and commented first. Relevant constructs and ideas from Cognitive Grammar (Langacker, 1987) are then borrowed for a tentative analysis of the forms for the past tense in English in terms of optimality.

III. ANALYSIS AND DISCUSSION

A. The Phenomenon of Speech Sounds

Language arises from and then serves for human interactions. Phonology is concerned with the description of how speech sounds are organized for that purpose. Any phonological model should then take the following basic facts of speech sounds seriously and try to incorporate them into it.

First, speech sounds are reference points (Rosch, 1975) of meaning. As the means to guide the hearer to what a

speaker intends, they have to be successfully parsed and interpreted by the hearer. This means that there must be pre-established form-meaning pairings in their minds for their references. A common mental representation of the association between a natural sound or kinetic melody (Sheets-Johnstone, 2003) and an entity, like the crying of a cuckoo bird, or that of the patting sound and the patting action, is clearly needed for the speech sounds "*cuckoo*" or "*pat*" to be initiated as a reference point. Let us call form-meaning associations at this level primary inter-subjectivity (Trevarthen et al, 2001).

For there to be utterance-meaning pairings, the members in a speech community have to interact. To say the human mind is hard-wired with a phonological system is relegating the very responsibility a researcher should undertake to God, let alone the circularity it incurs. If language is understood to have started by piecemeal, namely, from its proto-form to the fully-fledged one, the phonological system, along with its semantics, is necessarily emergent in nature (Beckner et al, 2009). The imitations of the crying sound of a cuckoo bird is sometimes for the hearer to look, sometimes for him or her to catch. Different orienting actions have different utterances, viz., different speech sounds. Sound-meaning associations at this level could be called secondary inter-subjectivity (Carpenter et al., 1998). The word "*cuckoo*" is an abstraction from these different usage events, hence emergent.

A third aspect of speech sounds is that they are produced by human mouths. This is a truism. But in what way is phonology related to the human mouth? There is a rather long tradition to disassociate phonetics, the study of speech sounds, from phonology, the patterning of them. Phonological analysis is carried out in terms of abstractions like features, phonemes, syllables etc. Scant attention has been given to the fact that all these constructs are meaningful only in reference to speech sounds, rather than alphabetic writing, which is only an approximation of speech. Real-time speech sounds are produced by physical movements in the mouth space along the time dimension. Spatial and temporal factors are part and parcel of speech sounds. Both phonetics and phonology shall not fail to see this.

Ideally, all the above factors should be taken into account to make any optimality-theoretic approach reasonable and compelling. In the following paragraphs, classic OT in phonology and its offshoots will be examined in terms of them.

B.   Classic (OT)

The following quote is a succinct summary of classic OT (Mohanan et al, 2010, p. 143-144):

Both (OT and Generative Grammar) theories share the assumption that Universal Grammar (UG) provides an inventory of universal distinctive features; OT also assumes that UG provides an inventory of universal constraints. The feature inventory defines the set of potential input segments in human languages that serve as phonemic, i.e., contrastive segments across languages, while the combination of feature inventory and constraints define the set of cross-linguistically possible sets of (i) phonetic segments, (ii) sequences of phonetic segments, and (iii) pairings of phonemic-phonetic segments/segment sequences. The features and constraints provided in UG are available to all individual grammars, hence it follows that all individual grammars contain the same set of distinctive features and constraints. Individual grammars vary in their interactions between constraints, formally expressed as ranking. Structural differences between individual grammars are due solely to differences in ranking. Constraints hold on phonetic representations (output constraints, i.e. markedness constraints), as well as on the pairing between phonemic and phonetic representations (input-output constraints, i.e. faithfulness constraints). There are no constraints on phonemic representations (Richness of the Base). Thus, all individual grammars have the same set of contrastive units (phonemic inventory) and there are no limits on sequences of phonemic segments. However, languages vary in the set of permissible phonetic segments and in the permissible sequencing of those segments. The assumption that UG provides universal inventories of features and constraints allows us to explain phonological patterns that appear recurrently across genetically unrelated languages while the assumption that individual grammars vary in the ranking of universal constraints allows us to explain the typological variations between such recurrent patterns.

Everything seems to be tidy and neat at first blush. Yet at a closer look it is not very difficult to spot unconvincing elements in its argumentation. Let's first examine the role UG plays in this theory. Suppose that there does exist such a thing as UG, which determines the inventories of both features and constraints. But it is just something mysterious and beyond reach. Practitioners still have to go to linguistic generalizations for its availability. The inventories they use to generate output forms are downright abstractions from linguistic data and have no relation whatsoever with UG. It then follows that the positing of UG is superfluous. It has no practical use for researchers.

Theoretically, however, UG is a powerful tool to make this theory a theory: it gives their feature and constraint inventories an absolute status to qualify them as a fiat that dictates its phonological modeling as if the phonological system is fully autonomous and allow not even the smallest bit of the penetration of human cognition. Could this arbitrary powerful tool be too powerful?

OT stipulates that input should be determined by the universal feature inventory. This amounts to saying that for any linguistic utterance, there are an infinite number of possible phonemic inputs, hence the Richness of the Base hypothesis. Is this compatible with the reality of speech sounds mentioned above? Both intuition and observation tell us that language works in terms of chunks like syllables or multi-syllables that are motivated by specific things or events in specific situations. It is true that they have to abide by the constraints exerted by the anatomy of the human mouth to include some phonological features, but the situated nature of an utterance makes it evident that not just any feature(s) goes. The sounds /kʊkuː/ of "*cuckoo*" and /pæt/ of "*pat*" are clearly motivated by natural sounds and speech sounds other than those in the slashes are less likely candidates for them. Phonemic inputs can never run as wild as the

Richness of the Base hypothesis suggests. The hypothesis of features as the sole determining factor of input is in this sense seriously flawed.

The claim that the set of phonemic segments is constraint-free is also not without its problems. Every phonemic segment in this set is said to have two values: unmarked and marked. Markedness serves the purpose of making distinctions. A natural corollary of this presumption is that markedness is an outcome of the functional pressure for making distinctions. That is to say, before markedness becomes universal, the input inventory has already gone through an optimizing process. It follows that markedness could have never been free of constraints. In this consideration, the feature inventory is redundant for this model as it is not what brings about markedness constraints.

Then what are the factors that have led to markedness? Given the discussion of the phenomenology of speech sounds in section II, a reasonable speculation will be that distinctive features are not the primitive force for inputs. An utterance, be it a syllable or a multi-syllable, can become the first word without having to be distinctive: there are simply no other words to be distinct from. Even when there are quite a number of words, they can still be different without resorting to binary features. The chances that a distinctive feature is invoked to distinguish meaning could only be found in minimal pairs or sets. But these occasions on which marked forms are purposefully exploited to distinguish meanings are rather rare if the whole vocabulary of a language is taken into account.

In light of the above reasoning, one wonders whether the concept "markedness" is a valid construct for phonological study at all, as it is useful only spasmodically. For phonemic segments, it is meaningful only when there are phonemic segments like, /□//𝒮ʃ/, /♦//♒/, /&//ℳ̥ɒ/, /♠//❆/, /✋//✋🖳/, /☾//✿/ or /✎//✑/ that are different by only one feature and can be paired. The degree of markedness among segments in cross relationship is very difficult to establish. Admittedly, there are languages, e.g. English, that make use of such contrastive pairs profusely. However, there are also ones, e.g. Chinese, in which these kinds of distinctions are only marginally exploited. Markedness thus cannot be applied to the relationships of all phonemic segments.

For phonemic sequencing, markedness is chiefly concerned with syllable structure and complexity of onset and coda. Phonemic sequences as input are usually those with elements that are to be assimilated, changed, deleted, filled, or repeated. Most of these inputs are morphologically motivated except for the one in which the last segment is devoiced. As to the latter, there is clearly a phonetic factor involved, namely, the last segment is usually given less force in speech. Again, input is not constraint-free and the so-called markedness are functionally or phonetically motivated. Markedness as a cover term, consequently, can then be dispensed with, too.

If there is any situation in which markedness constraints are of relevance, it is where there is competition between the functional pressure to distinguish meaning and the principle of economy to tackle phonetic difficulties in the articulation of lexical as well as morpho-syntactic forms that they will try to maintain marked forms. But even this can be override by frequency of use, which will be shown in later discussions. By restricting constraints strictly within the domain of articulation, classic OT has stripped phonology of the chance to relate itself with two other domains that are also crucial for verbal communication, time and human cognition.

The linear ordering of linguistic communication exposes all segments in it to the regulation of timing. Naturally produced pre-linguistic vocalizations are all streamlined because they are produced with no conscious control. Once they become prototypical linguistic units through consensual interaction (Maturana, 1978), more linguistic items will develop on the basis of them by adding or subtracting a feature, creating minimal pairs. When they are combined to express more complex ideas, say, a proposition, timing will come to play its part. Where there are segments at either side of the lexical boundaries that do not go with each other, the speaker is faced with a choice: either to smooth them up or not, depending on the time allowed for its articulation.

Whatever form, once finding its way into linguistic communication, is registered along with its context by human memory. Memorized chunks constitute a part of the reservoir of inputs for future uses. The forms entered into memory are subject to complicated process of mental operations like paradigmatic associations of ready-made chunks and further syntagmatic divisions of them, the results of which, phonemic segments, lexical substrates, neighborhood relations of lexical items and contextual clues, to name just a few, will all have to bear upon later linguistic productions. If the base is really rich, it should have included all the above factors.

In view of the inadequacy in the Richness of the Base hypothesis and the looseness in the definition of markedness, one cannot help but wondering whether the core of OT, the interaction of markedness and faithfulness constraints, can ever successfully make the predictions as it has been claimed to be able to.

As a summary of the above analysis, Classic OT is faced with at least the following three problems. The first is the logical fallacy in the claim that both the feature and the phonemic inventories are universal and the former determines the latter. It is baffling not only because it allows one universal to determine another, but also because it dissembles the circularity that is inherent in the model as a whole. Different languages in the world use different phonemic inventories. Any phonological theory should account for why and how they have come into being. The account given by classic OT is that they are the result of the competition between markedness and faithfulness constraints. When asked where the constraints are from, its answer is that they are innate. But in truth, they are typological generalizations (Kager, 1999) from existing phonemic inventories, which is what it is supposed to explain from the very beginning.

The second problem concerns the bivalent nature of constraints. As argued above, markness arises out of functional pressures and presupposes the existence of other linguistic chunks. Because a Markedness constraint is bivalent, it

either works on the input or not. Once it is put to work, a changed form will be the result; if not, the input is considered to be faithful. In this way a ranking of markedness and faithfulness is established. The model then predicts that whenever such an input occurs, it obeys this ranking of constraints. Is this really the case? It is not. For example, an explanation for the concatenations of "*going to*" and "*want to*" in English into "*gonna*" and "*wanna*" respectively could indeed be contrived in terms of the interactions of markedness and faithful constraints. But at the same time it will rule out the possibility that the former could still appear in formal communications. It allows no variations in hyper- and hypo-corrections, and as a result, leads to wrong predictions.

The co-existence of the original input and its altered form also shows that faithfulness constraints in classic OT is too restrictive. It is true that the phonological system is to some extent autonomous, viz., it has its own rules of organization when pressed by time, which is why some plausible markedness constraints can be derived. But human cognition, especially its memory and categorizing abilities, can organize and manipulate it for pragmatic purposes once it is acquired. Faithfulness can work at another dimension—speakers' volition.

Lastly, the assumption that speech sounds are conditioned only by markedness and faithfulness constraints has made linguistic communication too static a phenomenon. Time is another dimension on which linguistic communication unfolds itself. It may press the speaker to concatenate linguistic segments when it is limited; while it may allow full substantiation of them when it is ample. The ironing out of markedness or its maintenance, viz, being faithful, is a function of the time allowed for a specific utterance. In the events that phonological changes do take place, e.g. "*gonna*" and "*wanna*" as mentioned above, they are likely to become routines and speakers then have two forms of the same input to be faithful to. The ranking of markedness and faithfulness constraints provided by classic OT may be useful for inter-linguistic differentiation of languages, but intra-linguistically it leads to wrong predictions.

C. Functional OT

Functional OT is based on the traditional principle that the tension between clarity and ease is the driving force of phonology. That is, constraints originate from speakers' phonetic knowledge (Donegan et al, 1979). "The source of markedness constraints as components of grammar is this knowledge. The effect phonetic knowledge has on the typology of the world's sound systems stems from the fact that certain basic conditions governing speech perception and production are necessarily shared by all languages, experienced by all speakers, and implicitly known by all. This shared knowledge leads learners to postulate independently similar constraints. The activity of similar constraints is a source of systematic similarities among grammars and generates a structured phonological typology" (Hayes et al, 2004, p. 1-2).

Such a line of reasoning epitomizes in Boersma and colleagues (1999), where they argue that: 1) the initial state of grammar consists of a generating mechanism for speech sounds with variable degrees of difficulty and a mechanism for the detection and differentiation of them; 2) the phonetic mechanism gradually develop into grammatically meaningful constraints with relative distinctive capabilities; 3) phonetic functions determines the development of the ranking of constraints.

For Boersma, the constraints defined by classic OT are false. Its descriptions for the effects of markedness constraints can be derived from three functional principles: minimization of effort, maximization of recognition and minimization of categorization. Basing on these principles, he successfully derives different obstruent systems (Boersma, 2003).

What he proves is that phonological grammar is not abstract knowledge, but the result of the interaction of the generating mechanism of speech sounds and the mechanism for speech perception. Such a model has provided a much more reasonable account of the emergence of phonological systems without resorting to a hypothesis of innateness that is neither provable nor falsifiable. But it does not explain why and how the obstruent system, or any other ones like the plosive, the labial, etc. should rise at all.

D. The Combined Model of OT and Exemplary Theory (ET)

Exemplar theory was first developed as a model of similarity and classification in perception. It is later extended to be a model of perception, production, and the consequences of the perception-production loop over time. Its relevance for the study of phonology is as follows:

The exemplar approach associates with each category of the system a cloud of detailed perceptual memories. The memories are granularized as a function of the acuity of the perceptual system (and possibly as a function of additional factors). Frequency is not overtly encoded in the model. Instead, it is intrinsic to the cognitive representations for the categories. More frequent categories have more exemplars and more highly activated exemplars than less frequent categories (Pierrehumbert, 2001, p. 143).

Out of the concern with problems like phonological opacity, too many solutions in the input system and inability to cope with variation that are inherent in classic OT, and also the observation that all these problems might be accommodated by introducing exemplar theory into an OT framework, van de Weijer (2012) proposes to combine them "in such a way that all strengths will be retained and all weaknesses disappear" (p. 49). The combined model goes as follows:

In the combined theory, there is no infinite set of candidates, like in classical OT. Instead, the candidates are the exemplar tokens as in ET, so they are individual-specific and come with frequency-related information, which is a direct function of the degree of exposure to a particular item. The evaluation metric, i.e. the grammar, picks out one of these,

viz. the candidate that is best suited for use on that particular occasion (taking into account stylistic factors, for instance) (van de Weijer, 2009, p.123).

In this model, the innateness hypothesis is also discarded. Constraints are generalizations across the data to which language learners are exposed. We can see clearly such theoretical position is just the same as is held by functional OT, which is a compelling model for the generation of the phonemic inventory. Since this combined theory shares the same epistemological stand and aims at explaining how inter-word phonology is carried out with exemplary inputs, it is potentially an integral part of functional approach to phonological optimality.

But as a newly proposed model, there is still much to work on. One fundamental issue is what the production mechanism ought to be like. Should we posit a strategy of abstraction in speakers, as a result of which they hypothesize an underlying form on the basis of available tokens? Or should we simply mark a specific token as the underlying form, so that it has a special status among the exemplars and then the phonological grammar could change it into another one if the phonological environment gives occasion to assimilate (van de Weijer, 2009)? Such an issue is directly related to the dubious nature of the organization of exemplar clouds. Though it will take time and effort to come up with a solution, the combined model seems to be on the right track.

E. Cognitive Grammar

As has been discussed previously, speech production involves articulatory gestures (Browman et al, 1992) produced on the temporal scale, with human cognition as a very attentive manager. What is produced always leaves a trace on the mind. Specific gestures, along with their combinatory values for other gestures, are represented in the forms of motor routines as well as speech sounds in the human mind. The form of a specific sound given rise to by a specific gesture, when juxtaposed with other sounds, will invariably take on slight or dramatic changes, depending on the time allowed for the articulation of the involved gestures. The changed sound relative to the original one is not likely to be reinvented every time because human memory may well have retained it. In such a case, the specific way that has led to the changed sound is also registered, becoming a potential attractor of similar cases. Exemplar clouds formed in this way are thus hereditary.

After long enough time, a cloud tends to stabilize. But with gestures, which are dynamic by nature, pulling the strings behind the scenes, one can never say it is absolutely static. New exemplars will continue to come in. This is why language is forever on the drift. Neither can one say a cloud is chaotic. Every bit of speech sounds in it is a useful construction, so long as the speech flow is meaningfully parsed.

Though an optimality-theoretic framework in phonology that adopts phonological constructions as basic units of analysis is by no means clear at present, it could be argued that, since "the major empirical motivation for Construction Grammar is the need to develop a system of grammatical description in which the marked constructions (more or less 'idiom-like' forms of expressions) are represented in the same formal system as the regular, 'core' patterns or rules" (Kay, 2002, p.1), and both the functional model by Boersa and the combined model by van de Weijer are pointing to a functional approach to phonological optimality, such a model is in no way far-fetched. The permission of the co-presence of constructions and their derived ones in cognitive grammar (Langacker, 1987) renders it very much in compliance with both functional models and will help fix a production mechanism for the combined model.

A preliminary delimitation for such a model is that the only units permitted are constructions, which may include phonemes, consonant clusters, syllable structures, stress and rhythmic patterns etc. that constitutes an inheritance hierarchy. Also included are the second-order schemas that are abstracted from the above-listed constructions. Such a hierarchy is the result of two broad kinds of constraints: time-conditioned articulatory difficulty and pragmatic aptness which not necessarily works on an either-or basis. To give a pre-taste of how this model could work, a hierarchy of past tense formation in English is presented in the following graph.
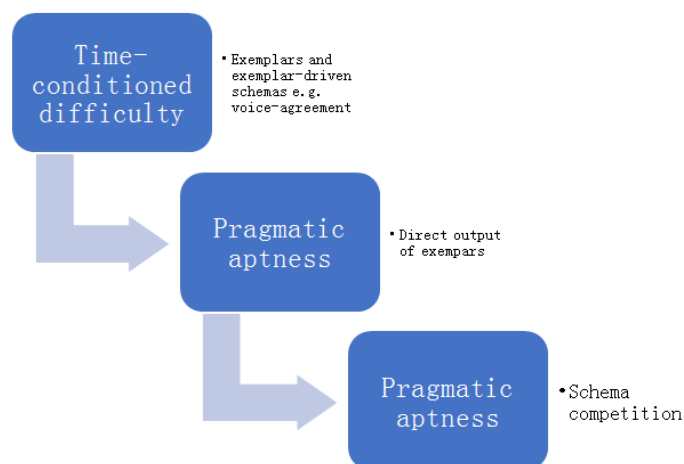
What find their ways into the input box at the top is superordinate concatenative schemas like [–ed] → [voiceless] / _

[voiceless], [–ed] → [voiced] / _ [voiced] and superordinate ablaut schema like [i] → [a] that are immanent in concrete past forms like *"walked"*, *"bragged"* and *"drank "*respectively. Because the past forms of most verbs have been committed to memory, they could be directly reproduced as output in the production stage as indicated by the second box. In cases where the past form of a verb is not registered, abstract schemas would come into play. Speakers would improvise a past form for the verb according to the extent to which these schemas are entrenched. Note that the two output stages indicated by the middle and bottom boxes are not two different stages, they are diagrammed so for the purpose of illustration. In actuality, they are the same process.

Sketchy as it, such a model can already predict what has caused variations, how varied forms are opted for, and why errors that are recorded in Berko's famous experiments (Berko, 1958) can happen. If such a model is fully substantiated at this level, it could be as competitive as any existing one. In consideration that it also has the potential to include the origins of different past forms, e.g. how strong and weak forms for the past tense have come into being, it is even a better one as it could incorporate diachrony as well as synchrony, and allows gradience.

IV. CONCLUSION

Language is messy, speech sounds may be even more so. But messy parts of language and speech sounds are still legitimate parts of language and for this reason they are also legitimate objects of scientific investigation. Theorists like generative linguists and classic OT practitioners more often than not try to make a distinction between the core and the peripheral to validate their hypothesis that language is a tidy autonomous system. But this tidy-mindedness does not lead to tidy theories.

Language and its phonological module are complex adaptive systems. Every piece of them results as an optimum from communicative pressures of different natures. The pursuit for the explanation of them in terms of functional optimality represents a psychologically plausible paradigm and will surely bear fruits.

REFERENCES

[1]   Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Schoenemann, T. (2009). Language is a complex adaptive system. Position paper. *Language Learning, 59 Supplement* 1: 1-26.
[2]   Berko, J. (1958). The child's learning of English morphology. *Word*, 14(3-4), 150-177.
[3]   Blevins, J. (1997). Rules in Optimality Theory: two case studies. In Roca, I. (ed.), *Derivations and constraints in phonology*, 227-260. Oxford: Clarendon Press.
[4]   Boersma, P. (1998). Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives. The Hague: Holland Academic Graphics.
[5]   Boersma, P. (2003). The odds of eternal optimization in OT. In Holt E. (ed.) *Optimality Theory and Language Change*, 31-65. Dordrecht: Kluwer.
[6]   Browman, C. P., & Goldstein, L. F. (1992). Articulatory phonology: an overview. *Phonetica*, 49, 155–180.
[7]   Carpenter M, Nagell K, Tomasello M. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development 63(4)*: 1–166.
[8]   Donegan, P. & Stampe. D. (1979). The study of natural phonology. In Dinnsen D., (ed.), *Current Approaches to Phonological Theory*, 126-173. Bloomington, IN: Indiana University Press.
[9]   Harnad, S. (1990). The symbol grounding problem. *Physica D,* 42: 335-346.
[10] Hayes, B. and Steriade D. (2004). Introduction: The phonetic bases of phonological markedness. In Hayes B., Kirchner R. and Steriade D. (eds.), *Phonetically Based Phonology*, 1-33. Cambridge: Cambridge University Press.
[11] Kager, R. (1999). Optimality Theory. Cambridge: Cambridge University Press.
[12] Kay, P. (2002). An informal sketch of a formal architecture for construction grammar. *Grammars*, 5(1), 1-19.
[13] Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive Science* 10, 1–40.
[14] Maturana, H. (1978). Biology of Language: The Epistemology of Reality. In Miller, G. A. & Lenneberg E. (Eds.), *Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg*, 27-63. New York: Academic Press.
[15] Mohanan, K. P., Archangeli D., & Pulleyblank D. (2010). The emergence of Optimality Theory. In Uyechi L. & Wee L. H. (ed.), *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, 143–158. Stanford University: Center for the Study of Language and Information.
[16] Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In by Bybee J. & Hopper P. (ed.) *Frequency and the emergence of linguistic structure*, 137-57, Amsterdam: John Benjamins.
[17] Prince, A. & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. RuCCS-TR-2, Rutgers Center for Cognitive Science. New Brunswick: Rutgers University. ROA-XYZ.
[18] Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology,* 7, 532-547.
[19] Sheets-Johnstone, M., (2003). Kinaesthetic memory. *Theoria et Historia Scientiarum*, 7 (1): 69-92.
[20] Trevarthen C. & Aitken K. J. (2001). Infant intersubjectivity: research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry,* 42, 3-48
[21] van de Weijer, J. M. (2009). Optimality theory and exemplar theory. *Phonological Studies* 12: 117-124.
[22] van de Weijer, J.M., (2012). Grammar as Selection: Combining Optimality Theory and Exemplar Theory. Kougaku Shuppan, Nagoya.

Zhenjun Song was born in Lianyungang, China in 1970. He received his Master degree in linguistics from Shanghai International Studies University, China in 2007.

He is currently a Ph.D. student in Shanghai International Studies University, China. His research interests include phonetics, phonology and morphology.