Saudi Journal of Humanities and Social Sciences (SJHSS) **ISSN 2415-6256 (Print)**
Scholars Middle East Publishers **ISSN 2415-6248 (Online)**
Dubai, United Arab Emirates
Website: http://scholarsmepub.com/

An Economic Model of Foreign Language Learning and Its Implications for Econometric Analysis

Yanqing Jiang*, Xianru Ge
School of Economics and Finance, Shanghai International Studies University Shanghai, China

***Corresponding author**
Yanqing Jiang

Article History
Received: 04.12.2017
Accepted: 11.12.2017
Published: 30.12.2017

DOI:
10.21276/sjhss.2017.2.12.5

Abstract: In this paper, we build a formal, rigorous economic model for foreign language learning, taking language input and language output as two crucial factors working jointly to generate language competence. The central idea behind our model is that the learner, constrained by limited resources, is confronted with a tradeoff between input-oriented training and output-oriented training. First, one major contribution of our modeling lies in its effort to reconcile the Input Hypothesis with the Output Hypothesis in the literature by showing that one important aspect of the learner's learning strategy, i.e. the input-output mix, is crucially dependent on the learning environment. Second, our modeling provides guidelines for data based empirical research, where the partial effects of the variables of interest can be estimated and hypotheses on the directions and magnitudes of the effects can be tested, econometrically. Third, our modeling also provides insights that can assist language instructors in making and fine-tuning teaching strategies.
Keywords: Economic, Foreign Language, learning, knowledge.

INTRODUCTION

There have been so many "competing" theories in foreign (second) language learning (acquisition) [1]. Though different theories may emphasize different influencing factors underlying the process of foreign language learning, no theory can deny the crucial role that language input plays in the process.

Input in the second language literature refers to the language to which a learner is exposed either orally or visually [1]. In fact, no model of second language acquisition does not avail itself of input in trying to explain how learners create second language grammars [1]. Language learning cannot happen in vacuum, and it is an incontrovertible fact that some sort of input is essential for language learning. What is controversial is the type and perhaps amount of input necessary for second language development and what additional information might also be necessary for the development of second language knowledge.

A framework within which input plays a dominant role is the input hypothesis, developed by Krashen [2, 3]. In his view, acquisition takes place by means of a learner's access to comprehensible input. That is, only a certain portion of the input is useful for the development of linguistic knowledge: the input at an *i+1* level, or a little bit beyond the learner's current system. Krashen argues that language acquisition depends on comprehensible input and in the classroom the instructor's main role is to provide the learner with suitable listening and reading materials so that the learner can receive comprehensible input.

However, many researchers later challenged the Input Hypothesis by supplying abundant evidence supporting that comprehensible input alone, a necessary condition as it is, is not sufficient for language acquisition [4-7]. They argue the ability to understand meaning conveyed by input does not automatically transform into the ability to use a linguistic system to express meaning [4, 8]. In order for learning to occur, language production or output is needed because production involves syntactic and grammatical processing. The idea that output or language use could be part of the learning mechanism itself was not seriously considered prior to Swain's [9] important paper, in which she introduced the notion of comprehensible or "pushed" output. In fact, the impetus for Swain's original study was the lack of second language development by immersion children even after years of academic study in that second language. Swain studied children learning French in an immersion context, suggesting that what was lacking in their development as native-like

[1] In the current study, we use "foreign language learning" consistently throughout the text, though much of our modeling and discussion here applies to "second language learning/acquisition" or "foreign language acquisition" too.

speakers of French was the opportunity to use language productively as opposed to using language merely for comprehension. As Swain [4] stated, "output may stimulate learners to move from the semantic, open-ended, non-deterministic, strategic processing prevalent in comprehension to complete grammatical processing needed for accurate production. Gass and Selinker [10] provided a summary of the four functions of output based on Swain's ideas (known in the literature as the Output Hypothesis [9, 6, 5]: testing hypotheses about the structures and meanings of the target language, receiving crucial feedback for the verification of the formed hypotheses, promoting a shift from a more semantic mode of processing to a more syntactic mode, and developing fluency and automaticity in interlanguage production.

In the output-driven hypothesis, Wen [11] asserts second language learning with output can lead to better outcomes than learning without output. Wen looks on output as the driving force for language learning as well as an eventual learning outcome itself. Encouraged to try out a productive activity, learners are more likely to notice what they lack in performing the assigned task and to be driven to learn what they want. Input, nevertheless, functions as an enabler that provides support for the learner to carry out productive activities. Addressing the order of pedagogical activities, the output-driven hypothesis reverses the order of learning by putting output before input to serve as a driving force for L2 learning.

In the meta-analysis of studies that compared comprehension-based instruction (CBI) and production-based instruction (PBI), Shintani *et al.*, [12] find that CBI is more effective for converting input into intake at the initial stage of acquisition while PBI may assist the process of accessing the partially acquired knowledge because production involves deeper processing. These findings accord with theoretical claims on the role of input and output by VanPatten [13] and Swain [4]. VanPatten [13] suggests that learners' limited processing capacity may prevent learners from converting input to intake, and that forcing learner to produce can interfere with their ability to attend linguistic form. Therefore, at the initial stage, providing sufficient input is the first priority, and certain manipulation of the input may be conducive to the conversion of input to intake. On the other hand, production, or output, is needed at a later stage to consolidate what has been partially acquired and to invite new input.

To date, there still exists a lack of consensus regarding the relative importance of input and output in the process of foreign language learning, and it seems that the respective relative importance of the two is contingent on many influencing variables concerning the learner, the learning environment, and the interaction between the learner and the learning environment. Seeing this point, can we make a reasonable attempt to reconcile the ideas of the Input Hypothesis with those of the Output Hypothesis and place them together into the same unified analytic framework? To this end, we build a formal, rigorous model for foreign language learning, taking input and output as two crucial factors that work jointly to generate language competence. We hope that, by exploiting the model's rigorous mathematical derivations, we can extend our insights into areas beyond the reach of our intuition and informal reasoning. In our model, we focus on the critical issue of the learner's resource allocation in the decision process of foreign language learning.

The rest of this paper is organized as follows. In Section 2, we present a formal, rigorous economic model for foreign language learning. In Section 3, we discuss the appropriateness of the model, focusing on its implications for econometric analysis. Section 4 presents more thoughts on and further modeling of foreign language learning based on the analyses and discussions in the previous sections. Finally, Section 5 concludes this paper.

The Model

In this section we build a formal model, which is an economic model presented mathematically, centering on the important issue of resource allocation in the decision process of foreign language learning. On a heuristic level, the decision process of foreign language learning, that is, the process of how a language learner, subject to her mental and resource constraints, considers the numerous tradeoffs involved in the process and makes a decision aiming to optimize her language learning outcomes can be thought of as being completely analogous to a production decision routinely studied in the discipline of economics. Therefore, in this section, we borrow heavily from economic theory and present a rigorous mathematical "production optimization" model for foreign language learning. The basic idea behind our model is that a language learner is a rational decision maker who, facing the tradeoff between receiving training by language input and receiving training by language output, makes an optimal decision as to the appropriate allocation of her scarce time and mental resources in order to achieve maximized (improvement in) language competence.

In our model, we define Y_{iT} as a stock variable denoting the accumulated level of language competence at time T for any representative language learner i. Mathematically, we can write Y_{iT} as

$$Y_{iT} = \sum_{s=0}^{T-1} \Delta Y_{is} \tag{1}$$

in which the zero point in time refers to the starting time point of the learning process and the total length of the time period $[0, T]$ is partitioned into evenly spaced time slots ΔY_{i0}, ΔY_{i1}, …, $\Delta Y_{i,T-1}$. The reason for the need to divide the entire time extension into small intervals is that by doing so we allow for the possibility that different intervals may involve potentially different values of the model's parameters. This is a very reasonable possibility because many factors underlying the efficiency and speed of language learning such as the learner's motivation and her mental and resource constraints may vary (substantially) across different time intervals. We define the change (i.e. increase or improvement, which is a flow quantity by definition) in accumulated language competence during any single time interval as

$$\Delta Y_{is} \equiv Y_{i,s+1} - Y_{is}, \quad s = 0, 1, \ldots, T-1, \tag{2}$$

where Y_{is} and $Y_{i,s+1}$ (stock values) denote the learner's level of accumulated language competence at time s and time $s+1$ respectively, and ΔY_{is} (a flow value by construction) denotes the change (i.e. increase or improvement) in the learner's level of language competence achieved during the time slot $[s, s+1]$.

Taking any representative time slot $[t, t+1]$ (for a specific language learner i), we can model her "production function" during this time slot as

$$\Delta Y_{it} = F(Q_i, a_{it}\Delta I_{it}, b_{it}\Delta O_{it}) \tag{3}$$

where $F(\bullet,\bullet,\bullet)$ denotes the functional relationship between the improvement in language competence (the dependent variable) and language training by input and output (independent variables), where ΔI_{it} and ΔO_{it} are the two main arguments of this production function (analogous to two production factors in a regular production function studied in economics), respectively measuring the amounts of training by input and training by output (flow quantities by definition) occurring during the time slot $[t, t+1]$ (and accordingly, notations such as I_{it} and O_{it} should be stock variables denoting the accumulated amounts of language training by input and by output received by learner i by the specific point in time t). As the quality of input and output can vary greatly in different situations, we have to control for the quality of them when measuring their quantities. Therefore, the quantities of ΔI_{it} and ΔO_{it} used in (3) should refer to normalized standard-quality quantities of training by input and by output that are measured in some constant-quality units. The coefficients a_{it} and b_{it}, which enter the function multiplicatively with ΔI_{it} and ΔO_{it}, are meant to be two efficiency parameters that are associated with the flows of standard-quality input and output respectively, indicating the efficiency rates associated with the transformation of language training by input and output into achieved (improvement in) language competence. Generally, the products $a_{it}\Delta I_{it}$ and $b_{it}\Delta O_{it}$ can be understood as the quantities of effective input and effective output processed by learner i during the time interval $[t, t+1]$. The third argument in the function, Q_i, which need not have a time subscript, is designed to be a person-specific, time-invariant variable aiming to capture a host of time-constant factors such as language learner i's innate ability (steady language aptitude, etc.), which can be assumed to vary across different learners but remain (virtually) unchanged for any specific learner over time.

The functional form $F(\bullet,\bullet,\bullet)$ in (3) is just like a black box in which input and output are allowed to interact with the learner and with each other via the learner in certain ways so as to jointly produce (improvement in) language competence in the learner. Therefore, our modeling accords with the interaction approach in the language learning literature, which subsumes certain aspects of the Input Hypothesis [2, 3] and the Output Hypothesis [9, 4, 5][2]. The efficiency coefficients a_{it} and b_{it} in our model take account of the roles played by input, interaction and output. On the one hand, input ΔI_{it} is absolutely important in language learning and no theory can overlook its importance. Input feeds an innate system to promote its growth [14]. The product $a_{it}\Delta I_{it}$ in our model can be thought of as the part of input that is effectively internalized (through internal processing and interaction) by the learner, which can be roughly called "intake" [15]. Therefore, the coefficient a_{it} attempts to measure the learner's ability to internalize what she has learnt or experienced from input to build her language competence. On the other hand, output is also intrinsically important for language learning. Output should be properly viewed as an integral part of the learning mechanism in itself, instead of being merely a way of producing what has already been learnt [10]. Output plays an important role in stimulating the learner to shift from the meaning-based, open-ended, non-deterministic, strategic processing prevalent in comprehension

[2] For more discussions on this issue, see Gass and Mackey (2015).

to a more syntactic mode of processing needed for accurate production [4]. In our model, the coefficient b_{it}, therefore, attempts to capture how efficiently the learner can utilize output (again, through internal processing and interaction) to develop syntax and morphology and form a complete grammatical processing mode necessary for native-like language production.

Since the production function in (3) is static by nature (with all its independent and dependent variables pertaining to the same time interval $[t,\ t+1]$), maximizing Y_{iT} in (1) reduces to the static problem of maximizing ΔY_{is} in each time slot $[s,\ s+1]$. Therefore, we face the problem in which the language learner, who is actually a decision maker and who is now conveniently assumed to be a rational one, maximizes her "production" (i.e. the increase in her level of accumulated language competence) in a representative time slot $[t,\ t+1]$, subject to her "cost constraint" in that time slot. The production maximization problem can then be modeled as

$$\max \quad F(Q_i, a_{it}\Delta I_{it}, b_{it}\Delta O_{it}), \text{ subject to } r_{it}\Delta I_{it} + w_{it}\Delta O_{it} \le \Delta M_{it} \tag{4}$$

Where the choice variables in the objective function $F(\bullet,\bullet,\bullet)$ are (standard-quality) input and output, ΔI_{it} and ΔO_{it}; the decision maker (the language learner) chooses the values of ΔI_{it} and ΔO_{it} optimally. By "optimally" we mean that the language learner chooses the values of ΔI_{it} and ΔO_{it} to maximize the objective function given her cost constraint as expressed in $r_{it}\Delta I_{it} + w_{it}\Delta O_{it} \le \Delta M_{it}$. If the objective function $F(\bullet,\bullet,\bullet)$ is strictly monotonic, the cost constraint will hold with equality: $r_{it}\Delta I_{it} + w_{it}\Delta O_{it} = \Delta M_{it}$. In the cost constraint, ΔM_{it} denotes the total "cost" that the language learner can afford to "pay" in order for an improvement in language competence, ΔY_{it}, to take place during the time interval $[t,\ t+1]$. This cost pertains to the total training time as well as mental resources (attention, cognitive energy, mental power used for overcoming anxiety, etc.) devoted to language learning during $[t,\ t+1]$ (and accordingly, the notation M_{it} should denote the accumulated level of total "cost" at time t, which is a stock variable). Later on, to fix ideas and to make practical issues much easier, we can think of ΔM_{it} as simply being the proportion (share) of total training time (the time when the learner's language learning process is "on") in the total length of the one-unit time interval $[t,\ t+1]$. The coefficients r_{it} and w_{it}, entering the cost constraint multiplicatively with ΔI_{it} and ΔO_{it}, refer to the quantities of time and mental resources that must be consumed, respectively, for receiving one standard-quality unit of training by input and one standard-quality unit of training by output. These can be regarded as the "prices" (in a broader sense) faced by this specific learner for obtaining one (standard-quality) unit of input and one (standard-quality) unit of output, respectively, during the time interval $[t,\ t+1]$. Later on, to make things simpler (but with no loss of generality), we can conveniently think of r_{it} and w_{it} as representing the mere time costs of obtaining one standard-quality unit of input and one standard-quality unit of output respectively.

To solve the optimization problem described in (4) above, a surefire way is to form the associated Lagrangian function, which is

$$L(\Delta I, \Delta O, \lambda) \equiv F(Q, a\Delta I, b\Delta O) + \lambda(\Delta M - r\Delta I + w\Delta O) \tag{5}$$

where the subscripts are dropped to avoid cluttering the notation. Assuming the objective function is mathematically well behaved and an interior solution exists, the first-order Lagrangian conditions are written as

$$\frac{\partial L}{\partial \Delta I} = \frac{\partial F}{\partial \Delta I} - \lambda r = 0 \tag{6a}$$

$$\frac{\partial L}{\partial \Delta O} = \frac{\partial F}{\partial \Delta O} - \lambda w = 0 \tag{6b}$$

$$\frac{\partial L}{\partial \lambda} = \Delta M - r\Delta I + w\Delta O = 0 \tag{6c}$$

Therefore, we can solve the equations for the values of the model's endogenous variables ΔI^* and ΔO^* that optimize the Lagrangian (as well as the objective function F in (4)) as functions of all the exogenous variables r, w, ΔM, a, b and Q.

If we take a specific (and very reasonable) functional form, i.e. the well-known Cobb-Douglas (CD) form, instead of the unspecified functional form in (33), we then have

$$\Delta Y_{it} = Q_i (a_{it} \Delta I_{it})^{\alpha} (b_{it} \Delta O_{it})^{\beta} \qquad (7)$$

with $0 < \alpha < 1$ and $0 < \beta < 1$. First, the CD functional form in (7) presumes that input and output are both necessary for language learning to take place; either zero input or zero output will result in zero ΔY_{it} according to the CD functional form. Second, the CD function in (7) is strictly increasing in its three arguments and exhibits "diminishing marginal returns" to both input and output. Mathematically, the first-order partial derivatives of ΔY_{it} with respect to ΔI_{it} and ΔO_{it} are positive while the second-order partial derivatives are negative:

$$\frac{\partial \Delta Y_{it}}{\partial \Delta I_{it}} = \alpha a_{it} Q_i (a_{it} \Delta I_{it})^{\alpha - 1} (b_{it} \Delta O_{it})^{\beta} > 0,$$

$$\frac{\partial \Delta Y_{it}}{\partial \Delta O_{it}} = \beta b_{it} Q_i (a_{it} \Delta I_{it})^{\alpha} (b_{it} \Delta O_{it})^{\beta - 1} > 0,$$

$$\frac{\partial^2 \Delta Y_{it}}{\partial \Delta I_{it}^2} = \alpha (\alpha - 1) a_{it}^2 Q_i (a_{it} \Delta I_{it})^{\alpha - 2} (b_{it} \Delta O_{it})^{\beta} < 0 \qquad \text{(since } \alpha - 1 < 0)$$

$$\frac{\partial \Delta^2 Y_{it}}{\partial \Delta O_{it}^2} = \beta (\beta - 1) b_{it}^2 Q_i (a_{it} \Delta I_{it})^{\alpha} (b_{it} \Delta O_{it})^{\beta - 2} < 0 \qquad \text{(since } \beta - 1 < 0)$$

Moreover, according to (7), whether language learning exhibits increasing, constant, or decreasing "returns to scale" depends on whether the case $\alpha + \beta > 1$, $\alpha + \beta = 1$ or $\alpha + \beta < 1$ holds. By increasing, constant, or decreasing returns to scale, we mean that if we increase both the quantities of ΔI_{it} and ΔO_{it} by K (any positive real number) times, then as a result, the quantity of ΔY_{it} will be increased by more than K times, exactly K times, or less than K times, respectively. Later in the next section, we will have a brief discussion of how to empirically test for returns to scale.

The optimization problem in (4) now becomes

$$\max\ Q_i (a_{it} \Delta I_{it})^{\alpha} (b_{it} \Delta O_{it})^{\beta}, \text{ subject to } r_{it} \Delta I_{it} + w_{it} \Delta O_{it} \le \Delta M_{it} \qquad (8)$$

The associated Lagrangian becomes (with the subscripts dropped)

$$L(\Delta I, \Delta O, \lambda) \equiv Q (a \Delta I)^{\alpha} (b \Delta O)^{\beta} + \lambda (\Delta M - r \Delta I + w \Delta O) \qquad (9)$$

The first-order conditions are, accordingly

$$\frac{\partial L}{\partial \Delta I} = Q \alpha (a \Delta I)^{\alpha - 1} a (b \Delta O)^{\beta} - \lambda r = 0 \qquad (10a)$$

$$\frac{\partial L}{\partial \Delta O} = Q (a \Delta I)^{\alpha} \beta (b \Delta O)^{\beta - 1} b - \lambda w = 0 \qquad (10b)$$

$$\frac{\partial L}{\partial \lambda} = \Delta M - r \Delta I + w \Delta O = 0 \qquad (10c)$$

where the first two of the three conditions (10a) and (10b) jointly yield

$$\frac{\alpha \Delta O}{\beta \Delta I} = \frac{r}{w} \tag{11}$$

which, together with the third condition (10c), leads to the solution values of optimal input and output (in standard-quality units):

$$\Delta I^{*} = \frac{\alpha}{\alpha + \beta} \frac{\Delta M}{r} \tag{12a}$$

$$\Delta O^{*} = \frac{\beta}{\alpha + \beta} \frac{\Delta M}{w} \tag{12b}$$

Inserting the values in (12a) and (12b) back into the CD production function in (7) (and putting back the dropped subscripts), the optimized ΔY_{it} can then be written as the following function (analogous to an inverse cost function in economics):

$$\Delta Y_{it} = \overline{A} Q_i \frac{a_{it}^{\alpha} b_{it}^{\beta}}{r_{it}^{\alpha} w_{it}^{\beta}} \Delta M_{it}^{\alpha + \beta} \tag{13}$$

With $\overline{A}$ being a constant defined as $\overline{A} \equiv \frac{\alpha^{\alpha} \beta^{\beta}}{(\alpha + \beta)^{(\alpha + \beta)}}$. Therefore, according to (1), language learning that can be achieved during $[0, T]$ (for individual i) is determined by

$$Y_{iT} = \overline{A} Q_i \sum_{t=0}^{T-1} \frac{a_{it}^{\alpha} b_{it}^{\beta}}{r_{it}^{\alpha} w_{it}^{\beta}} \Delta M_{it}^{\alpha + \beta} \tag{14}$$

Taking logs on both sides of (13) leads to the following linear relationship that can later serve as a basis for a linear panel data regression specification:

$$\ln \Delta Y_{it} = \ln \overline{A} + \ln Q_i + \alpha \ln(a_{it} / r_{it}) + \beta \ln(b_{it} / w_{it}) + (\alpha + \beta) \ln \Delta M_{it} \tag{15}$$

We will come back to equation (15) later in the next section for a discussion of its implications for econometric modeling and empirical testing.

An alternative frequently used functional form is the constant-elasticity-of-substitution (CES) function, which, in the current case, can be written as

$$\Delta Y_{it} = Q_i [(a_{it} \Delta I_{it})^{\rho} + (b_{it} \Delta O_{it})^{\rho}]^{1/\rho} \tag{16}$$

Where $0 \neq \rho < 1$.[3] It is easy to show that the CES function form is strictly increasing in its three arguments, mathematically,

$$\frac{\partial \Delta Y_{it}}{\partial \Delta I_{it}} = Q_i [(a_{it} \Delta I_{it})^{\rho} + (b_{it} \Delta O_{it})^{\rho}]^{(1/\rho)-1} a_{it}^{\rho} \Delta I_{it}^{\rho - 1} > 0 ,$$

[3] It is obvious that the CES functional form does not presume that language input and language output should both be necessary in the process of language learning, as the CD function does.

$$\frac{\partial \Delta Y_{it}}{\partial \Delta O_{it}} = Q_i[(a_{it}\Delta I_{it})^{\rho} + (b_{it}\Delta O_{it})^{\rho}]^{(1/\rho)-1} b_{it}^{\rho}\Delta O_{it}^{\rho-1} > 0$$

Further it can also be shown that the CES function is strictly quasi-concave. The optimization problem now becomes

$$\max \ Q_i[(a_{it}\Delta I_{it})^{\rho} + (b_{it}\Delta O_{it})^{\rho}]^{1/\rho}, \text{ subject to } r_{it}\Delta I_{it} + w_{it}\Delta O_{it} \le \Delta M_{it} \tag{17}$$

The associated Lagrangian now becomes (again with the subscripts dropped for less cluttering)

$$L(\Delta I, \Delta O, \lambda) \equiv Q[(a\Delta I)^{\rho} + (b\Delta O)^{\rho}]^{1/\rho} + \lambda(\Delta M - r\Delta I + w\Delta O) \tag{18}$$

The first-order conditions are

$$\frac{\partial L}{\partial \Delta I} = Q[(a\Delta I)^{\rho} + (b\Delta O)^{\rho}]^{(1/\rho)-1} a^{\rho}\Delta I^{\rho-1} - \lambda r = 0 \tag{19a}$$

$$\frac{\partial L}{\partial \Delta O} = Q[(a\Delta I)^{\rho} + (b\Delta O)^{\rho}]^{(1/\rho)-1} b^{\rho}\Delta O^{\rho-1} - \lambda w = 0 \tag{19b}$$

$$\frac{\partial L}{\partial \lambda} = \Delta M - r\Delta I + w\Delta O = 0 \tag{19c}$$

The first two of the three conditions (19a) and (19b) jointly lead to

$$\frac{a^{\rho}\Delta I^{\rho-1}}{b^{\rho}\Delta O^{\rho-1}} = \frac{r}{w} \tag{20}$$

which, combined with the third condition (19c), generates the solution values of optimal input and output (in standard-quality units):

$$\Delta I^{*} = \frac{b^{k} r^{k-1}}{b^{k} r^{k} + a^{k} w^{k}}\Delta M \tag{21a}$$

$$\Delta O^{*} = \frac{a^{k} w^{k-1}}{b^{k} r^{k} + a^{k} w^{k}}\Delta M \tag{21b}$$

where we have defined $k \equiv \rho/(\rho-1)$. Inserting the values in (21a) and (21b) back into the CES production function in (16) (and putting back the dropped subscripts), after a bit of rearrangement, the optimized ΔY_{it} can then be written as

$$\Delta Y_{it} = Q_i[(r_{it}/a_{it})^{k} + (w_{it}/b_{it})^{k}]^{-1/k}\Delta M_{it} \tag{22}$$

Based on (22), the summation in (1) implies that language learning achieved during $[0, T]$ (for individual i) can be expressed by

$$Y_{iT} = Q_i \sum_{t=0}^{T-1} [(r_{it}/a_{it})^{k} + (w_{it}/b_{it})^{k}]^{-1/k}\Delta M_{it} \tag{23}$$

The multiplicative form of the right-hand side of (22) becomes an additive form after taking logs on both sides (with a little rearrangement):

$$\ln \Delta Y_{it} = \ln Q_i - (1/k)\ln[(a_{it}/r_{it})^{-k} + (b_{it}/w_{it})^{-k}] + \ln \Delta M_{it} \tag{24}$$

In passing, it should be noted that by construction, $k \equiv \rho /(\rho - 1)$ and $0 \neq \rho < 1$ so that the sign of k depends on whether $0 < \rho < 1$ (which implies $k < 0$) or $\rho < 0$ (which implies $k > 0$) holds. In the next section, we will come back to equation (24) for a discussion of its implications for econometric modeling and empirical testing.

DISCUSSIONS

The two functional specifications presented in the preceding section, different as they are, do share a similarity in that they lead to the same major results, which can be summarized as follows:

1. The increased language competence ΔY_{it} is positively related to Q_i.
2. ΔY_{it} is also positively related to ΔM_{it}.
3. More importantly, ΔY_{it} is positively related to a_{it} / r_{it} and b_{it} / w_{it}.
4. The results above can be seen from both equations (15) and (24) though the two equations present very different specifications of a linear relationship. Further, from equations (12a), (12b), (21a) and (21b), we can find that:
5. Both the chosen optimal values of ΔI_{it} and ΔO_{it} are positively related to ΔM_{it}.
6. According to (12a) and (12b), the chosen optimal values of ΔI_{it} and ΔO_{it} are negatively related to their own "price", r_{it} and w_{it}, respectively.
7. In comparison with (v) above, according to (21a) and (21b), the chosen optimal values of ΔI_{it} and ΔO_{it} are negatively related to each other's "price", w_{it} and r_{it}, respectively, when $\rho < 0$ (or equivalently $k > 0$), and are positively related to each other's "price", w_{it} and r_{it}, respectively, when $0 < \rho < 1$ (or equivalently $k < 0$).

Intuitively, the exogenous variables of the models, i.e. r_{it}, w_{it}, ΔM_{it}, a_{it}, b_{it} and Q_i, have their meaningful interpretations. The "price" variables, r_{it} and w_{it}, generally represent the amounts of time and mental resources that must be "paid" for obtaining one standard-quality unit of training by input and one standard-quality unit of training by output, respectively, during the time interval $[t, t+1]$. To fix ideas and make things simpler (but with no loss of generality), we can conveniently consider r_{it} and w_{it} to be the mere time costs of obtaining one standard-quality unit of input and one standard-quality unit of output respectively. We hope that the values of r_{it} and w_{it} can serve as indexes or indicators reflecting the level of friendliness or effectiveness (etc.) of the learning environment during $[t, t+1]$. In sum, we (reasonably) assume the levels of r_{it} and w_{it} to be pertaining to environmental factors of language learning. For example, access to the Internet, say, newly available to an underdeveloped, non-native-speaking country would be expected to greatly reduce the level of r_{it} by exposing the learner to a much larger bulk of high-quality reading materials for language input.

The total "cost" ΔM_{it} is designed to represent total training time and mental resources (such as attention, cognitive energy, etc.) that are available to be devoted to language learning during $[t, t+1]$ Also, to fix ideas and simplify the issue, we can consider ΔM_{it} to be simply the ratio of the length of total training time (the time when the learning process is "on") to the total length of the (one-unit) time interval $[t, t+1]$[4]. We hope that the level of ΔM_{it} can serve as an indicator or proxy variable that reflects the level of interest, motivation or diligence (etc.) of and inside the learner during the time interval $[t, t+1]$. In sum, we use ΔM_{it} to capture underlying factors related to intrinsic characteristics of the learner, such as her interest, motivation, diligence, etc.

The efficiency coefficients a_{it} and b_{it} are associated with the efficiency in the transformation of (flows of) input and output into achieved (improvement in) language competence. Therefore, $a_{it}\Delta I_{it}$ and $b_{it}\Delta O_{it}$ represent the amounts of effective training by input and effective training by output received by the learner during the time interval $[t, t+1]$. The two coefficients a_{it} and b_{it}, which vary over time and across different learners (so that they are treated as variables), capture the absorption capacities of the learner with respect to training by input and output. The absorption capacities are in turn associated with underlying factors such as the learner's cognitive ability and learning techniques.

[4] As the denominator, i.e. the length of the time interval $[t, t+1]$ is unity by construction, the ratio of the length of total training time to the total length of $[t, t+1]$ equals the length of the total training time, which is the numerator.

Finally, the variable Q_i, which lacks a time subscript, is a person-specific, time-invariant variable that aims to capture time-constant factors such as the learner's innate ability (e.g. innate language aptitude, which is assumed to remain fixed over time). In comparison with the efficiency coefficients a_{it} and b_{it} that capture the learner's absorption capacities, Q_i refers to constant innate ability that is fixed for any individual but is heterogeneous across different individuals. Therefore, in our models above, Q_i is assumed to vary across different learners but remain (virtually) unchanged for any specific learner over time.

With the exogenous variables bearing their own interpretations, we can make use of the two log-linear equations (15) and (24) to design our empirical models for econometric analysis. Equation (15), which is based on a CD functional form of the production function, can be used to formulate a regression equation studying the returns to scale as well as the direction and magnitude of the partial effect of each exogenous variable mentioned earlier. The $\ln Q_i$ term in (15) can be transformed into an unobserved individual heterogeneity term that can be canceled out in, for instance, a fixed-effects (FE) or first-differencing (FD) panel data setting. Hypothesis testing with respect to the function's structural parameters, α and β, can be used for examining whether the functional relationship exhibits increasing, constant, or decreasing returns to scale. If the null hypothesis H_0: $\alpha + \beta = 1$ cannot be rejected, we can then reasonably assume that the functional relationship exhibits constant returns to scale, meaning that if both the quantities of ΔI_{it} and ΔO_{it} are increased by K (any positive real number) times, then the quantity of ΔY_{it} will be increased by exactly K times as a result. If instead the null hypothesis H_0: $\alpha + \beta = 1$ is rejected in favor of the alternative hypothesis H_1: $\alpha + \beta > 1$ (or $\alpha + \beta < 1$), this is evidence supporting increasing (or decreasing) returns to scale, which means that if both the quantities of ΔI_{it} and ΔO_{it} are increased by K times, the quantity of ΔY_{it} will be increased by more (or less) than K times as a result. If constant returns to scale is the case, then according to (15) there will be a one-to-one relationship between $\ln \Delta M_{it}$ and $\ln \Delta Y_{it}$, meaning that a 1% change in ΔM_{it} will lead to exactly a 1% change in ΔY_{it}.

Likewise, to construct a regression model based on equation (24), the $\ln Q_i$ term in the latter can also be transformed into an unobserved individual heterogeneity term that can be eliminated in, say, a fixed-effects (FE) or first-differencing (FD) panel data setting. As in (24) the coefficient on $\ln \Delta M_{it}$ is unity by construction, we can see that in this case, just like in the prior case, a 1% change in ΔM_{it} leads to exactly a 1% change in ΔY_{it}. However, the major problem with using equation (24) as the basis for our empirical specification lies in the nuisance parameter involved $\ln[(a_{it}/r_{it})^{-k} + (b_{it}/w_{it})^{-k}]$, which renders the method of linear regression infeasible.

More Thoughts and Further Modeling

Our discussions above lead us to believe that empirical results may vary in crucial ways in response to the theoretical model as well as the underlying assumptions we employ. To push things further and see what happens under the most generalized framework, we introduce the "cost function" for language learning and the associated "Hicksian demands" for input and output[5].

The cost function for language learning can be defined in the following way. First, we construct the cost minimization problem (dropping the subscripts hereinafter for cleaner notations):

$$\min(r\Delta I + w\Delta O), \text{ subject to } F(Q, a\Delta I, b\Delta O) \geq \Delta Y \tag{25}$$

where this time r, w, ΔY, a, b and Q are the exogenous variables, ΔI and ΔO are the choice variables, and we minimize the total cost incurred by input and output, subject to the condition that the resulted improvement in language competence is no less than ΔY. The associated Lagrangian function can be formed in a straightforward way and the related first-order conditions can be derived. We can then solve the first-order equations for the optimal values of ΔI and ΔO, which are denoted by ΔI^* and ΔO^*. Then the cost function for language learning is defined as

$$\Delta E = r\Delta I^* + w\Delta O^* \tag{26}$$

[5] An interested reader can refer to, for example, Jehle and Reny (2011) for more details of the theoretical background.

where ΔI^* and ΔO^* are in turn functions of the exogenous variables r, w, ΔY, a, b and Q, which are a version of the so-called "Hicksian demands" used in microeconomics. To simplify the notations and move further, we temporarily treat Q as a constant and define $x_1 \equiv a\Delta I$, $x_2 \equiv b\Delta O$, $p_1 \equiv r/a$ and $p_2 \equiv w/b$. Then the cost minimization problem can be rewritten as

$$\min(\ p_1 x_1 + p_2 x_2), \text{ subject to } f(x_1, x_2) \geq \Delta Y \tag{27}$$

The Hisksian demands associated with optimal values x_1^* and x_2^* can then be written as functions of p_1, p_2 and ΔY, i.e. $x_1^h(p_1, p_2, \Delta Y)$ and $x_2^h(p_1, p_2, \Delta Y)$, where the superscript *h* indicates explicitly that they are the Hicksian demand functions.

With the newly defined variables x_1, x_2, p_1 and p_2 (and treating Q as a constant), the production maximization problem in (4) can be recast as

$$\max\ f(x_1, x_2), \text{ subject to } p_1 x_1 + p_2 x_2 \leq \Delta M_{it} \tag{28}$$

The (optimal) values of x_1 and x_2 that solve the maximization problem in (28) are the so-called Marshallian demand functions, whose arguments are the exogenous variables p_1, p_2 and ΔM, i.e. $x_1^m(p_1, p_2, \Delta M)$ and $x_2^m(p_1, p_2, \Delta M)$, where the superscript *m* reminds us that these are Marshallian demand functions.

Regarding the relationship between the Marshallian demand functions and the Hicksian demand functions, a remarkable theorem in economics, the Slutsky equation, tells us that

$$\frac{\partial x_i^m(p_1, p_2, \Delta M)}{\partial p_j} = \frac{\partial x_i^h(p_1, p_2, \Delta Y^*)}{\partial p_j} - x_j^m(p_1, p_2, \Delta M)\frac{\partial x_i^m(p_1, p_2, \Delta M)}{\partial \Delta M} \tag{29}$$

with *i*, *j* = 1, 2, and ΔY^* being the maximized level of $\Delta Y = f(x_1, x_2)$ that can be achieved given p_1, p_2 and ΔM under (28) above. In the special case $i = j$, the Slutsky equation in (29) decomposes a price effect (which is the total effect) into two separate effects, i.e. the substitution effect and the income effect. The term on the left-hand side of (29) represents the total effect, the first term on the right-hand side of the equation captures the substitution effect, and the second term on the right-hand side pertains to the income effect. As in this case the substitution effect can always be shown to be negative, the magnitude of the total effect is then the sum of the magnitudes of the substitution effect and the income effect provided the latter is positive (Note the minus sign before it). That is to say in the case $i = j$, whenever the substitution effect is negative and the income effect is positive, the total effect (or alternatively called the price effect) is negative.

The substitution effect refers to the (hypothetical) change in the chosen values of x_1 and x_2 that would occur if p_1/p_2 (which can be called the "normalized relative price") were to change to it new level but the level of ΔY were kept the same as before. The income effect is defined as whatever is left of the total effect after the substitution effect is netted out. To fix ideas, consider a scenario where p_1 is lowered for some reason. As p_1 is defined as $p_1 \equiv r/a$, p_1 is lowered whenever r is lowered or (and) a is raised. Imagine that, ceteris paribus, r is lowered because newly available access to the Internet (say, in an underdeveloped, non-native-speaking country) is now exposing the learner to a much larger pool of high-quality reading materials as potential input. Then how will $x_1^m(p_1, p_2, \Delta M)$ respond (where $x_1 \equiv a\Delta I$ as defined above)? Will the learner choose to have more training by input and less training by output, or vice versa? The effect of the decrease in r (hence a decrease in p_1) is the total price effect (the total effect or the price effect, for short). The decomposition in the Slutsky equation implies that, in order to see the direction of the total price effect, we have to determine the direction of the income effect and compare its magnitude against that of the substitution effect. The direction of the substitution effect is always negative, meaning that whenever p_1 decreases (and p_1/p_2 decreases as a result, ceteris paribus), the learner, in response, would demand more input and less output if she were to (hypothetically) keep her ΔY at its original level, and by keeping her ΔY unchanged, she effectively leaves a portion of her ΔM unspent. The income effect is the effect on the change of the learner's chosen values of x_1 and x_2 when the portion of ΔM left unspent by the substitution effect is now spent.

In the scenario above, three possible situations may exist:

1. (Language input is a normal good). If it turns out that the income effect leads the learner to increase x_1, then a decrease in p_1 has a positive income effect as well as a positive substitution effect. In this case the magnitude of the total price effect is the sum of the magnitudes of the substitution effect and the income effect. Both the substitution effect and the income effect push x_1 larger when p_1 is lowered.
2. (Language input is an inferior good). If, alternatively, the income effect leads the learner to decrease x_1, then a decrease in p_1 has a negative income effect and a positive substitution effect. In this case, usually the substitution effect is the dominant effect so that the total price effect is positive. When p_1 is lowered, the substitution effect pushes x_1 larger but the income effect pushes x_1 smaller, and the net result is that x_1 becomes larger.
3. (Language input is a Giffen good). In very rare cases, the income effect leads the learner to decrease x_1 as in (ii) but this income effect is so strong that it dominates the substitution effect. In such cases the total price effect becomes negative. The net result is that a decrease in p_1 leads to a decrease in x_1.

Obviously, the analysis of the effect of a change in p_2 on the (optimally) chosen value of x_2 is completely analogous to those discussions regarding p_1 and x_1 above. Through all the discussions in the current and previous sections, we reach a few important conclusions, which we present in the next section of this paper[6].

CONCLUDING REMARKS

In this paper, we construct a formal, rigorous economic model for foreign language learning, taking language input and language output as two crucial factors that work interactively to generate the learner's language competence. We hope that, by resorting to the rigor of the model's mathematical derivations, we can arrive at important findings that are beyond the reach of experience and informal reasoning. We focus on the critical issue of resource allocation in the decision process of foreign language learning. The central idea underlying our model is that the language learner, being constrained by her limited resources that can be expended on learning, is necessarily confronted with the tradeoff between input-oriented training and output-oriented training. Seeing this, in any given time period, a rational learner would choose the respective amounts of training by input and training by output in order to seek to maximize her progress in language competence, subject to her resource constraints.

The contributions of this paper to the literature are threefold. First, the most prominent contribution of our modeling in this paper lies in its attempt to reconcile the Input Hypothesis with the Output Hypothesis by showing that one important aspect of the learner's learning strategy, i.e. her input-output mix, is heavily dependent on her learning environment. For instance, if the learner chooses to use a (relatively) large amount of input-oriented training compared with output-oriented training, it is not necessarily because input is more important than output in the process of language learning, but because in her learning environment, input is (relatively) "inexpensive" compared with output. As is already shown in (6a) and (6b) earlier, the learner would seek to choose the respective levels of input and output such that the ratio of the marginal return to input ($\partial F / \partial \Delta I$) to the marginal return to output ($\partial F / \partial \Delta O$) exactly equals the ratio of the "price" of input (r) to the "price" of output (w)[7] Second, our modeling in this paper provides guidelines for data based empirical research. When assuming a suitable functional form, a tractable regression model can be derived. We can then obtain relevant data, work on the regression model, and estimate the respective partial effects of the explanatory variables. Various hypotheses on the directions and magnitudes of the partial effects can be formulated and tested. Hypotheses on returns to scale can also be tested too. Third, our modeling in this paper also provides insights that can help language instructors to form better teaching strategies. In some cases the language learner may not have full information about her functional form, so that she is likely to miscalculate the optimal levels of the choice variables in her production function (This is an example of bounded rationality). In other cases, some psychological/behavioral mechanisms inside the learner (e.g. mental accounting, decision making under risk and uncertainty, the sunk-cost fallacy, the status quo bias and default options, projection biases, naïve procrastination, committing) may set in to distort the

[6] See Jiang (2000, 2001) for more discussions on the role of input in language learning. Also, see the Appendix of Jiang (2000) for a simplified graphical illustration of part of the modeling in the current and previous sections.

[7] Dividing both sides of (6a) by the corresponding sides of (6b) after moving the second term in each equation to the right-hand side yields the result.

learner's rational choice[8]. In these latter cases, one job the language instructor (who sometimes have a better knowledge of the learner's production function) can do is to help the learner deal with her "behavioral anomalies" and choose the optimal learning strategy (e.g. the most suitable expansion path of the input-output mix), based on the guidelines our modeling in this paper can provide.

REFERENCES

1. Gass, S. (1997). Input, Interaction, and the Second Language Learner, Lawrence Erlbaum, Mahwah, NJ.
2. Krashen, S. (1982). Principles and Practices in Second Language Acquisition, Pergamon, Oxford.
3. Krashen, S. (1985). The Input Hypothesis: Issues and Complications, Longman, London.
4. Swain, M. (1995). "Three Functions of Output in Second Language Learning," in G. Cook and B. Seidlhofer (eds), Principle and Practice in Applied Linguistics, Oxford University Press, Oxford.
5. Swain, M. (2005). "The Output Hypothesis: Theory and Research," in E. Hinkel (ed), Handbook on Research in Second Language Learning and Teaching, Lawrence Erlbaum, Mahwah, NJ.
6. Harley, B., & Swain, M. (1984). "The Interlanguage of Immersion Students and Its Implications for the Second Language Teaching," in A. Davies, C. Criper and A. Howatt (eds), Interlanguage (pp 291–311), Edinburgh University Press, Edinburgh.
7. Harley, B., & Hart, D. (1997). "Language Attitude and Language Proficiency in Classroom Learners of Different Starting Ages. *Studies in Second Language Acquisitions, 19,* 379–400.
8. Crookes, G. (1991). "Second Language Speech Production Research: A Methodologically Oriented Review. *Studies in Second Language Acquisition, 13,* 113–132.
9. Swain, M. (1985). "Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development," in S. Gass and C. Madden (eds), Input in Second Language Acquisition, Newburry, Rowley, MA.
10. Gass, S. M., & Slinker, L. (2001). Second Language Acquisition: An Introductory Course, 2nd edition, Lawrence Erlbaum Associates, Inc., New Jersey.
11. Qiufang, W. (2016). "The Production-Oriented Approach to Teaching University Students English in China. *Language Teaching, 1*, 1-15.
12. Shintani, N., Li, S., & Ellis, R. (2013). "Comprehension-Based Versus Production-Based Grammar Instruction: A Meta-Analysis of Comparative Studies. *Language Learning, 63*, 296-329.
13. VanPatten, B. (2004). "Input Processing in Second Language Acquisition," in B. VanPatten (ed), Processing Instruction: Theory, research, and commentary (pp. 5–32), Lawrence Erlbaum, Mahwah, NJ.
14. Schwartz, B. D. (1993). "On Explicit and Negative Data Effecting and Affecting Competence and Linguistic Behavior, *Studies in Second Language Acquisition, 15,* 147–163.
15. Corder, S. P. (1967). "The Significance of Learner's Errors, International Review of Applied Linguistics.

[8] The study of the various psychological/behavioral mechanisms leads us to realm of behavioral economics (see, for example, Just 2014); even a cursory discussion of these mechanisms is beyond the scope of the current paper.